

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 17:46:09 ; Search time 130 Seconds
(Without alignments)
5451.325 Million cell updates/sec

Title: US-09-830-810A-1
Perfect score: 1277
Sequence: 1 aagcgcgagcgagcgcgagga.....acaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCUTS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	54	4.2	5467 1	US-07-745-206A-12 Sequence 12, Appl
2	54	4.2	5467 2	US-08-311-363-12 Sequence 12, Appl
3	53.2	4.2	7218 1	US-08-232-463-14 Sequence 14, Appl
4	50.8	4.0	6232 3	US-08-456-2008-11 Sequence 11, Appl
5	49.6	3.9	400 4	US-09-056-556-179 Sequence 179, App
6	49.6	3.9	400 4	US-09-072-596-174 Sequence 174, App
7	49.6	3.9	400 4	US-09-072-596-174 Sequence 179, App
8	49	3.8	7175 1	US-08-455-543A-8 Sequence 8, Appl
9	49	3.8	7175 2	US-08-193-0788-8 Sequence 8, Appl
10	49	3.8	7175 2	US-08-223-305C-8 Sequence 8, Appl
11	49	3.8	7175 2	US-08-149-097D-8 Sequence 8, Appl
12	49	3.8	7175 3	US-08-949-386-8 Sequence 8, Appl
13	49	3.8	7175 3	US-08-450-562-8 Sequence 8, Appl
14	49	3.8	7175 4	US-08-964-709A-8 Sequence 8, Appl
15	49	3.8	7175 4	US-08-450-272-8 Sequence 8, Appl
16	49	3.8	7175 4	US-08-450-273-8 Sequence 8, Appl
17	49	3.8	7177 4	US-09-268-163-7 Sequence 7, Appl
18	49	3.8	7266 3	US-08-713-118-1 Sequence 1, Appl
19	49	3.8	7266 3	US-09-452-007-1 Sequence 1, Appl
20	49	3.8	7362 1	US-08-455-543A-7 Sequence 7, Appl
21	49	3.8	7362 2	US-08-193-0788-7 Sequence 7, Appl
22	49	3.8	7362 2	US-08-223-305C-7 Sequence 7, Appl
23	49	3.8	7362 2	US-08-149-097D-7 Sequence 7, Appl
24	49	3.8	7362 2	US-08-949-386-7 Sequence 7, Appl
25	49	3.8	7362 3	US-08-450-562-7 Sequence 7, Appl
26	49	3.8	7362 3	US-08-984-709A-7 Sequence 7, Appl
27	49	3.8	7362 4	US-08-450-272-7 Sequence 7, Appl

28	49	3.8	7362 4	US-08-450-273-7 Sequence 7, Appl
29	49	3.8	7364 4	US-09-268-163-5 Sequence 5, Appl
30	49	3.8	7376 4	US-09-268-163-3 Sequence 3, Appl
31	48	3.8	4403765 3	US-09-103-840A-2 Sequence 2, Appl
32	46.6	3.6	745 1	US-08-036-555B-163 Sequence 163, App
33	46.6	3.6	745 1	US-08-469-569-163 Sequence 163, App
34	46.6	3.6	745 1	US-08-249-322A-163 Sequence 163, App
35	46.6	3.6	745 1	US-08-469-526A-163 Sequence 163, App
36	46.6	3.6	745 2	US-08-734-591A-163 Sequence 163, App
37	46.6	3.6	745 2	US-08-469-660-163 Sequence 163, App
38	46.6	3.6	745 3	US-08-341-018-51 Sequence 51, Appl
39	46.6	3.6	745 3	US-08-470-335-163 Sequence 163, App
40	46.6	3.6	745 3	US-08-735-021-163 Sequence 163, App
41	46.6	3.6	745 3	US-08-734-664A-163 Sequence 163, App
42	46.6	3.6	745 3	US-08-470-339-163 Sequence 163, App
43	46.6	3.6	745 4	US-08-467-602-163 Sequence 163, App
44	46.6	3.6	745 5	PCT-US94-05083C-159 Sequence 159, App
45	46.6	3.6	745 5	PCT-US95-06846A-163 Sequence 163, App

ALIGNMENTS

RESULT 1
US-07-745-206A-12
Sequence 12, Application US/07745206A
Patent No. 5429921
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McQue, Rm
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: Human Calcium Channel Compositions and
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A
FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 5:1504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5467 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (Genomic)
FEATURES:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390..3392, 3396..3488, 3498..3538, 3543..3581, 3585..3587, 3591..3626, 3630..3689, 3693..3737, 3744..3746, 3750..4823, 4827..4841, 4845..5006, 5010..5096, 5100..5306, 5310..5366, 5370..5465)
US-07-745-206A-12

```

Query Match Similarity      4.2%;      Score 54;  DB 1;  Length 5467;
Best Local Similarity      46.7%;      Pred. No. 0.00037;
Matches 111;  Conservative 0;  Mismatches 195;  Indels 0;  Gaps 0;

QY      303  CGAGGCGCCGACAGCTGTGCAGAGCTGTGCAGAGTGTGCGGAGCCGAGCCGAGCCGATCGGGTTC 362
DB      2800  CCGCGCGCCGACCGGACGCCACAGAGAGAGAGCGCGCGGCGGCCCGCGAGAGCGCGAGAGCG 2855
QY      363  CTGTCAACCCCGTGTGTCAAGCGCGCGCGCGGAGATCCCGCGATCTTGACAGACCTTAC 422
DB      2860  CGAGCGAGGCCAGAGGCGCCCTGAGAGGCGCGCGGCGGACCAACCGCGCGGTCTCCCGAGAGA 2915
QY      423  CCCGTTCTTGTCGTCGTCGAGCCTTCTGTGAGCCTCTCTCTCTCACTGAGAGTTGCGGAGGCGAG 482
DB      2920  GCGCGGCGCGAGCGGAGAGCGCCCGACCGCCGACCGCGGACCGCGACCGCGAGCGCGGCGC 3039
QY      483  GCAGACACCCACGAAAGGAGAGAGGAGCCCGGACATCTTGCGGAGCCCGGAGAACCGAGCC 542
DB      2980  GTGCGCGCGCGCGCAAGAGGCGAGCGCGCGCGCGCACCGCGGCGGCGGCCCGCGAGCGGGGCG 3039
QY      543  GAGAGAGGTGGCGCGGAGAGAAAGCGATCCCCAGCGCGGAGAGCGGAGCGGATTTCA 602
DB      3040  CCGGAGAGCGCGAGAGCGGAGAGAGCCGCGCGGCGGACCGCGGCGCGCGCAACAGCGCGCA 3099
QY      603  GGCTGCAAGGCGAGCGCGGCTGTGGAACAGACGACACACCGAGAGACCGGAGACAGTGTGCG 662
DB      3100  GCGTCTCTCAAGAGGCTGTGTGAAGAGAGACACGAGAGAGAGGCGCAAGAGAGAGAGGCG 3155
QY      663  GCGCAT 668
DB      3160  TGAGAT 3165

RESULT 2
US-08-311-363-12
: Sequence 12. Application US/08311363
: Patent No. 5676958
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: Human Calcium Channel Compositions and
: TITLE OF INVENTION: Methods
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-2926
: COMPUTER READABLE FORM:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/311,363
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-51506
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619)238-0999
: TELEFAX: (619)238-0062
: INFORMATION FOR SEQ ID NO: 12:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 5467 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3349..3386, 3390..3392, 3396..3488, 3495..3539, 3543..3581, 3585..3587, 3591..3626, 3630..3689, 3693..3737, 3744..3746, 3750..4823, 4827..4841, 4845..5006, 5010..5096, 5100..5306, 5310..5366, 5370..5465)
US-08-311-363-12

Query Match      4.2%   Score 54; DB 2; Length 5467;
Best Local Similarity 46.7%; Pred. No. 0.00037;
Matches 171; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY    303 CGGGGCGCCGACGCTGCAGCCTTCGAGGATGCCGAGCCCCGACGCCGATCGGGTTC 362
DB    2800 CGCGGCGCCGACCGAGCGACGACAAGAGAGGCGCGGGGGCCCCCGAAGCGCGCAAGCGC 28559

QY    363 CTGTCAACCCCCTGTGCCACACCGCGCGCCGCGGAGATTCGCCGATCTTGGAACCGTAGC 422
DB    2860 CGGGCGGAGCGCCACGAGCCCCCGAGGGCGCGCCGCGCACACCGGCGCGGCTCTCCCGAGGA 2919

QY    423 CCCGTTCTCGCTCGACTTCTGTGGCGCTCTCTCTCTCACTGGAGGTGGGGAGGAG 482
DB    2920 GGCGCGCGAGGGGAGCGCCCGAACGCCACCGCGCACCGGACACGAGATCCGACAAAGA 29797

QY    483 GCACAACCCACGAGAAGAGAGGAGCCCGGGCATCTCGGGGACCAGGGAACCGGAGCC 542
DB    2980 GTGCGCGCGCGCCACGAGGCGGAGCGCGCGCGGACACCGGCGCGGCCCGCCGAGCGGGGCC 30399

QY    543 GAGAGAGGTGGCGCGCGAGAAAAGCGTCCCCCGACCGCGGAAAGCGAGGAGCGCATTTCA 602
DB    3040 CCGGAGAGCGGAGACCGGGAGAGACCGCGCGCGGCGACCGGGCCCGGACAAAGCGCA 30999

QY    603 GGCTGCAAGGCGCGCGGTGGAGGACGACACCAACCGAGGAGCCGGAACAGTGTGGC 662
DB    3100 GCGTGTCAAGGAGCTGTGAGAAAGAGACCAAGAGAAAGAGGCCACGGAAGAAAGAGGC 3159

QY    663 GCGCAT 668
DB    3160 TGAGAT 3165

RESULT 3
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHRIFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOXP1 VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

```


APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mark, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 179:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-967-179

Query Match 3.9%; Score 49.6; DB 4; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.0015;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 458 CCTACTGAGAGTTGGCGGAGGACGAGACACCCAGAGGAGGAGGCCCGGCAT 517
DB 119 CTTCACTCCAGAGCGCCGACGCGGCAACGCGGCAACGCGGCGTGGCGGCA 178
QY 518 CCTCGGGAGCCCGGGAACCGGAGCCGAGAGAGTGGCCGCGAGGAAACGCGTCCCCAGC 577
DB 179 CGGCGGAAACGCGCGGAAACGCGGCAACACACACACGCGCGCCGCGGACACAGG 238
QY 578 CGCGAAGGAGAGGAGGCGGATTTAGAGTGAAGGCGAGCGCGGTGGAGACAGCAGCAC 637
DB 239 CGGCGACGCGCGGCGCGCGCGGCGGGAACCGCGGAAACCGCGGAAACCGCGGCGAC 298
QY 638 CACCGAGAGACCGGAACAGTGTGCGCGGATGAGTGTGAGCTGAGAGGAGAGGCGCAT 697
DB 299 CGGACACCGCGCGGCAACAGGCAACGCGGCGCAACGCGGCGGACCGCGGCGGCGCAC 358
QY 698 GTCTGCGGACAGAGTGGCTCAGG 721
DB 359 CGGCGGCGACGCTGCACTCTCAGG 382

RESULT 8
US-08-455-543A-8
Sequence 8, Application US/08455543A
Patent No. 5792846
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven

APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: May 31, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-455-543A-8

Query Match 3.8%; Score 49; DB 1; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGCTGAGGTGAACCGCCCGGAGCGCTCGGTGAGTGTCACTCGGGCGC 309
DB 2751 GAGGCCCCGAAAGGCGAGAGCCGGGAGCCCGGTGCCCCGGAGAGCGCGCGCGCAC 2810
QY 310 CGCAGCTGAGGCTGAGGAGGTGCGAGCGAGCCCGGAGTCCGGGTCTCTGTCA 369
DB 2811 CGCA-GCCACAGCAAGAGAGCGCGGGGCC--CCGAGAGCGGAGCGCGCGCGA 2867
QY 370 CCCGTGGCCACCGCCCGCGGAGATCCCGCGATCTTGACAGCCGTAGCCCCGTTG 429
DB 2868 GGCCTCAGGCCCCGAGGGCGCGCGCGCACACCGCGCGGTCCCGGAGAGCGCGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGGCGCGGAGCGCGCACCGCGCACCGGACAGAGATCCAGAGAGGCGCC 2987
QY 490 CCCACGAGGAGAGGAGGAGCGCGCGCATCTCGGGGACCGGAGCCGAGCGAGAG 549
DB 2988 GGGCGCAAGGGGAGAGCGCGCGCGCGCGCACCGCGCGGCGCCCGGAGCGCGAG 3047
QY 550 GTGGCCCCGAGAAAGCGGTCCCGCACCGCGAGCGAGAGGCGGATGTTCAAGCTGCA 609
DB 3048 GCGGAGAGCGGGAGAGAGCGCGCGCGCGCACCGGCGCGCGCACAGGCGCGCTGCT 3107
QY 610 GGGCAGCGCGGCGGTGGAGGAGAGCGCACCGAGAGAGCGGACAGTGTGGCGCGCAT 668
DB 3108 CACGAGGCTGTGGAGAGAGAGCACAGAGAGAGGCGCACAGAGAGAGGCTGAGAT 3166

RESULT 9
US-08-193-078B-8
; Sequence 8, Application US/08193078B
; Patent No. 5846757
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
; STREET: 1660 UNION STREET
; CITY: SAN DIEGO
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,078B
; FILING DATE: 07-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-53607
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999

TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; NAME/KEY: 5' UTR
; LOCATION: 1..143
; NAME/KEY: 3' UTR
; LOCATION: 6855..7175
US-08-193-078B-8
Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGCTGAGGTGAACCGCCCGGAGCGCTCGGTGAGTGTCACTCGGGCGC 309
DB 2751 GAGGCCCCGAAAGGCGAGAGCCGGGAGCCCGGTGCCCCGGAGAGCGCGCGCAC 2810
QY 310 CGCAGCTGAGGCTGAGGAGGTGCGAGCGAGCCCGGAGTCCGGGTCTCTGTCA 369
DB 2811 CGCA-GCCACAGCAAGAGAGCGCGGGGCC--CCGAGAGCGGAGCGCGCGCGA 2867
QY 370 CCCGTGGCCACCGCCCGCGGAGATCCCGCGATCTTGACAGCCGTAGCCCCGTTG 429
DB 2868 GGCCTCAGGCCCCGAGGGCGCGCGCGCACACCGCGCGGTCCCGGAGAGCGCGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGGCGCGGAGAGCGCGCGCGCGCACCGGACAGAGATCCAGAGAGTGTGCGC 2987
QY 490 CCCACGAGGAGAGGAGGAGCGCGCGCATCTCTCGGGGACCGGAGCCGAGCGAGAG 549
DB 2988 GGGCGCAAGGGGAGAGCGCGCGCGCGCGCACCGCGCGGCGCCCGGAGCGCGAG 3047
QY 550 GTGGCCCCGAGAAAGCGGTCCCGCACCGCGAGCGAGAGGCGGATGTTCAAGCTGCA 609
DB 3048 GCGGAGAGCGGGAGAGAGCGCGCGCGCGCACCGGCGCGCGCACAGGCGCGCTGCT 3107
QY 610 GGGCAGCGCGGCGGTGGAGGAGAGCGCACCGAGAGAGCGGACAGTGTGGCGCGCAT 668
DB 3108 CACGAGGCTGTGGAGAGAGAGCACAGAGAGAGGCGCACAGAGAGAGGCTGAGAT 3166

RESULT 10
US-08-223-305C-8
; Sequence 8, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P5:9739)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-223-305C-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGGTGAGGAGCCGCGGAGCGCTCGGTGAGGTTCACCTCGGGCGC 309
DB 2751 GAGGCCCCGAAAGCGAGAGGGGAGCCCGGTGCGGAGAGCGCCCGCGCCAC 2810
QY 310 CGCAGCTGAGCTTGAGGAGTCCGAGCCGAGCCCGATCGGTTCTCTCA 369
DB 2811 CGCA-GCCAGAGCAAGAGGCGCGCGGCGC--CCCGAGCGCGAGCGCGCGCGA 2867
QY 370 CCCCCTGACCGCGCGCGCGGAGATCCCCGGGATCTTGGAGACCGTAGCCCGCTTC 429
DB 2868 GGCCTCAGCGCGCGCGCGCGCGGCGGACCGCGCGCGCTCCCGGAGGCGCGCC 2927
QY 430 TCGTCCGTGAGCTTCTGTGGGCTCTCTCTCACTGAGGTTGCGGAGGCAAGCAGACA 489
DB 2928 GAGCGGAGAGCCCGAGCGCAAGCGCGGCGGAGCGGAGATCGAGCAAGAGTGGCC 2987

QY 490 CCCAGAAAGGAGAGGGAGCCCGGCACTCTGGGAGACCCGGAGAACCGGAGCCGAGAG 549
DB 2988 GAGCGCAAGGGCGAGCGCGCGCGGCGGCAACCGCGCGGCCCCCGAGGGGGCGAG 3047
QY 550 GTGGCCGCGAAGAGCGGTCCCGAGCCCGGAGCGAGAGGGCGATGTTCAGGCTGCA 609
DB 3048 GCGAGAGCGGGAGAGAGCCGCGCGCGCGGCGGCGAGCGGCGCGGAGCGGAGCTGCT 3107
QY 610 GGGCAGCGCGGTGAGAGCAGCCAGCCAGCCGAGAGACCGGAAAGTGTGCGCGCAT 668
DB 3108 CACGAGGCTGTGAGAGAGAGCAGCCAGAGAGAGGCCACGAGAGAGGCTGAGAT 3166

RESULT 11
US-08-149-097D-8
Sequence 8, Application US/08149097D
Patent No. 5874236
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,097D
FILING DATE: 05-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,231
FILING DATE: 13-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899

FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-55038
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-149-097D-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGGTGAGTGAACCCGCGCCGCGAGCCTTGCTGCTCACTCGGCGC 309
DB 2751 GAGGCCCCCAAGCGGAGCGCGGAGCCGCTGCGCGGAGAGCGCGCGCGCGC 2810
QY 310 CGCAGCGCTGAGCGCTGCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGC 369
DB 2811 CGCA-GCCACAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2867
QY 370 CCCCGTGGCCAGCGCGCGCGCGCGAGATCCCGCGATCTTGCGACACCGCGCTTC 429
DB 2868 GCGCCAGC 2927
QY 430 TCGTCGTGACCTTCTGTGCGCTTCTCTCTCACTGAGGTTGCGGAGCGCGACACA 489
DB 2928 GAGCGGGAACCCCGACCGCACCGCGCGCACCGGACCAAGATCCAGCAAGAGTGC 2987
QY 490 CCACAGAGAGAGAGAGAGCGCGCGCGATCTCGGAGACCGGAGCGCGAGAGAG 549
DB 2988 GCGCGCAAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 3047
QY 550 GTGCGCCGAGAGAAAGCGGTCCCGACCGCGAGCGAGAGAGCGATGTTCAAGCTGCA 609
DB 3048 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCGCACCGCGCGCGCGCGCGCGCT 3107
QY 610 GGGCAGAGCGGAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 668
DB 3108 CACGAGGCTGTGAGAGAGAGCAACGAGAGAGAGCGCGCGCGCGCGCGCGCTGAGT 3166

RESULT 12
US-08-949-386-8
Sequence 8, Application US/08949386
Patent No. 6030623
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Alison
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,386
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/290,012
FILING DATE: 11-AUG-1994
APPLICATION NUMBER: 08/149,097
FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 519808
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-949-386-8

Query Match 3.8%; Score 49; DB 3; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGGTGAGTGAACCCGCGCCGCGAGCCTTGCTGCTCACTCGGCGC 309
DB 2751 GAGGCCCCCAAGCGGAGCGCGGAGCCGCTGCGCGGAGAGCGCGCGCGCGC 2810
QY 310 CGCAGCGCTGAGCGCTGCGAGCGGAGCGCGCGCGATCCCGCGATCGGCTTCTGTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2867
QY 370 CCCCGTGGCCAGCGCGCGCGCGAGATCCCGCGATCTTGCGACACCGTAGCCCGCTTC 429
DB 2868 GCGCCAGC 2927
QY 430 TCGTCGTGACCTTCTGTGCGCTTCTCTCTCACTGAGGTTGCGGAGCGCGACACA 489
DB 2928 GAGCGGAGCGCGAGAGAGCGCGCGCGCGCGCGCACAGATCCAGCAAGAGTGC 2987
QY 490 CCACAGAGAGAGAGAGAGCGCGCGCGATCTCGGAGACCGGAGCGCGAGAGAG 549
DB 2988 GCGCGCAAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 3047
QY 550 GTGCGCCGAGAGAAAGCGGTCCCGACCGCGAGCGAGAGAGCGATGTTCAAGCTGCA 609

Db 3108 CACGAGCTGTGAGAGAGAGACCAACGAGAGAGGCGCAGGAGAGAGGCTGAGAT 3166

RESULT 14

US-08-984-709A-8
Sequence 8, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Stauderman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,709A
FILING DATE: 02-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 587-5360
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-984-709A-8

Query Match 3.8%; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGGTGACAGTGAACCCGCGCCGACGCGCTGTGACAGTGTCTACTGGGGCC 309
Db 2751 GAGGCCCGGAGAGGCGGAGAGCGCGGAGCGCGGTGCGCGGAGAGCGCGCCGCGCAC 2810
QY 310 CGCAGCGTGCAGCTGTGAGGAGTGCAGCCGACCGCCGACCGCGATCGAGTTCTGTCAA 369
Db 2811 CGCA-GCAGCAGCAAGAGGCGCGCGGCGCC--CCGAGAGGCGCGGAGCGAGCGCGCGCA 2867
QY 370 CCCGCTGCGCAGCGCGCGCGCGGAGATCCCGCGGATCTTGCAGACCGTAGCCCGCTTC 429
Db 2868 GCGCCAGCGCGCGAGGCGCGCGCGGAGCAGCAGCGCGCGGCTCCCGAGAGAGCGCGCC 2927
QY 430 TCGTCGAGACTTGTGTGCGCTCTCTCTCACTGAGAGTTGGGAGGAGGAGGAGCA 489

Db 2928 GAGCGGAGCGCCCGACCCGACCGCGCGCACCGGACACAGATCCGAGCAAGAGTGGCC 2987

QY 490 CCCAGAGAGAGAGGAGGAGCGCGGATCCTCGGGAGACCGGGAGACCGGAGCGAGAG 549

Db 2988 GCGCCAGAGGCGAGCGAG 3047

QY 550 GTGCGCGCGAGAGAGGCGGTCCCGCAGCCGCGAGCGAGAGGCGGATGTTCAAGCTGCA 609

Db 3048 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3107

QY 610 GCGGAGCGCGCGGTGAGAGCAGCAGCCACCGGAGAGCGGAGACGTGGCGCGCAT 668

Db 3108 CACGAGCTGTGAGAGAGAGACCAACGAGAGAGGCGCAGGAGAGAGGCTGAGAT 3166

RESULT 15

US-08-450-272-8
Sequence 8, Application US/08450272
Patent No. 6387696
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Alison
APPLICANT: Felman, Daniel
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,272
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,350
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: 08/336,257
FILING DATE: 7-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083
FILING DATE: 28-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/311,363
FILING DATE: 23-SEPT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,012
FILING DATE: 11-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: 4-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,078
FILING DATE: 07-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,097

FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
TELEPHONE: (619) 238-0999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-272-8

Query Match 3.8%; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTCGGTGCAAGTGAACCCGCGCGGACGCTCGGTGCAAGTTCATCTCGGGCCG 309
DB 2751 GAGGCCCGGAGGCGGAGCGGAGCGGCGGAGCCCGGTGCCCCGAGGAGCGCGCGCGCAC 2810
QY 310 CGCAGCGTGCAGCTGAGGGTGCAGCGCCGCGGACGCGCCGAGTCCGAGTCCGAGTTCGTCA 369
DB 2811 CGCA-CGCAAGCAAGAGGCGCGCGGAGCC-CCGAGGCGCGAGCGAGCGCGCGCA 2867
QY 370 CCCCGTGGCCACCGCGCGCGCGGAGATCCCGCGCATCTTGCGCAGACCGTAAGCCCGGTTTC 429
DB 2868 GAGCCAGGCGCGGAGGCGCGCGCGGACACCGCGCGGCTCCCGAGAGAGGCGGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGCTCTCTCTCTCACTGAGGTTGCGGAGGCAAGCAGACA 489
DB 2928 GAGCGGAGCGCGGAGCGCGCGCGCGCATCCGAGATCCGAGCAAGAGTGCGCC 2987
QY 490 CCCACGAAAGGAGAGGAGCGCGGATCTCGGAGACCGGAGACCGGAGCGGAGAG 549
DB 2988 GCGCGCAAGGCGAGGCGCGCGCGGACCGCGCGGCGCCCGGAGCGGCGCGGAG 3047
QY 550 GTGGCCGCGAGAAAGCGGTCCCGCAGCGCGGAGCGAGGAGGCGATGTTCAAGCTGCA 609
DB 3048 GCGGAGAGCGGAGGAGCGCGCGCGCGCACCGGCGCGGCGCGGCGCGGCGCGAGCTGCT 3107

QY 610 GGGCAGCGCGGCTGGAGAGCAGCAGCCACCGGAGGACCGGAAAGTGTGCGCGCAT 668
DB 3108 CACGAGGCTGTGAGAGAGGAGCAACGAGAAAGAGGCCACGAGAGAGAGGCTGAGAT 3166
Search completed: April 6, 2004, 20:38:15
Job time: 138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 16:19:29 / Search time 3686 Seconds
(without alignments)
10345.633 Million cell updates/sec

Title: US-09-830-810a-1
Perfect score: 1277
Sequence: 1 aagggcggcgagggcgaggga.....acaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:*
1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pin:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_pig:*
27: em_gse_vr1:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	643	50.4	675	10	BB641267 BB641267
2	642.4	50.3	666	10	BB264222 BB264222
3	591.4	46.3	615	10	BC071693 BC071693
4	582.4	45.6	595	12	BC084538 BC084538

C 5	570.4	44.7	594	10	BF471866	BF471866	UI-M-BH3-
C 6	497.2	38.9	525	10	BB704019	BB704019	BB704019
C 7	473.8	37.1	499	9	AT854700	AT854700	UI-M-BH0-
C 8	448.4	35.1	491	10	BB703259	BB703259	BB703259
C 9	413.8	32.4	491	10	BB704648	BB704648	BB704648
C 10	405.8	31.8	436	10	AU044294	AU044294	AU044294
C 11	398.2	31.2	425	10	BB705931	BB705931	BB705931
C 12	396.4	31.0	590	9	AU023153	AU023153	UI-M-BH3-
C 13	393	30.8	450	10	BE946858	BE946858	UI-M-BH3-
C 14	392.2	30.7	419	10	BB704449	BB704449	BB704449
C 15	367.4	28.7	521	10	BB703869	BB703869	BB703869
C 16	367	28.8	427	10	BB706957	BB706957	BB706957
C 17	364.2	28.5	434	10	BB699732	BB699732	BB699732
C 18	347.2	27.2	391	10	BB704146	BB704146	BB704146
C 19	333.8	26.1	400	10	BB700620	BB700620	BB700620
C 20	319	25.0	905	28	B2114495	B2114495	CH230-421
C 21	316.2	24.9	441	10	BB704848	BB704848	BB704848
C 22	282.2	22.1	301	13	EX637052	EX637052	EX637052
C 23	264.6	20.7	315	10	BB248342	BB248342	BB248342
C 24	227.4	17.8	642	13	BY719704	BY719704	BY719704
C 25	224.8	17.6	654	28	A2225053	A2225053	R2C1-23-6
C 26	211.6	16.6	251	9	AV359580	AV359580	AV359580
C 27	208.6	16.3	251	9	AV358045	AV358045	AV358045
C 28	203.2	15.3	249	9	AV357766	AV357766	AV357766
C 29	203	15.9	290	10	BB063113	BB063113	BB063113
C 30	201.8	15.8	539	9	AL904756	AL904756	AL904756
C 31	200.8	15.7	909	13	EX755136	EX755136	EX755136
C 32	200	15.7	286	10	BB060381	BB060381	BB060381
C 33	199.8	15.6	246	10	BB060134	BB060134	BB060134
C 34	198	15.5	539	9	AL904892	AL904892	AL904892
C 35	197.8	15.5	642	12	BM681479	BM681479	BM681479
C 36	197	15.4	643	12	BU098465	BU098465	BU098465
C 37	196.8	15.4	526	9	AL904839	AL904839	AL904839
C 38	196.8	15.4	534	9	AL904820	AL904820	AL904820
C 39	196.8	15.4	559	9	AL904774	AL904774	AL904774
C 40	196.8	15.4	575	9	AL904855	AL904855	AL904855
C 41	196.4	15.4	576	9	AL904863	AL904863	AL904863
C 42	196.4	15.4	578	9	AL904817	AL904817	AL904817
C 43	196.4	15.4	579	9	AL904815	AL904815	AL904815
C 44	196.4	15.4	579	9	AL904816	AL904816	AL904816
C 45	196.4	15.4	604	9	AL904765	AL904765	AL904765

ALIGNMENTS

RESULT 1
LOCUS BB641267 675 bp mRNA linear EST 26-OCT-2001
DEFINITION BB641267 RIKEN full-length enriched, 10 days neonate cortex Mus
musculus cDNA clone AB30014H23 5', mRNA sequence.
ACCESSION BB641267
KEYWORDS BB641267.1 GI:16476392
SOURCE
ORGANISM Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 675)
Arkawa,T., Carinci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,O.,
Komoto,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arkawa,T., et al. 2001)

TITLE
JOURNAL
COMMENT
Unpublished (2001)
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

Wael, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.

Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Komoto, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamakura, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences *Mamm. Genome*. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
SOURCE

```

1..666
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A830014H23"
/tissue_type="cortex"
/dev_stage="10 days neonate"
/lab_host="DH108"
/clone_1fb="RIKEN full-length enriched, 10 days neonate cortex"
/note="Site 1: Sali; Site 2: Banhi; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGGAGAAGACATCCAAAGCTCTTTTITTTTTTTTN 3'. cDNA was prepared by using Trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGACTTCGAGTGTAATTATAATATCCCCCCCCC 3'. cDNA was cleaved with XhoI and BanHI. Vector: a modified plasmid pSC(+) after bulk excision from Lambda DLI 1."

```

ORIGIN

Query Match	50.3%	Score 642.4	DB 10	Length 666
Best Local Similarity	99.6%	Pred. No. 6.2e-87		
Matches 665	Conservative	0	Mismatches 1	Indels 2
Qy	588	GGAGGGCGATGTTCCAGCTGCAGGGCAGGCCGGGTGGAGCAGCAGCCACCAGGAGA	647	
Db	1	GGAGGGCGATGTTCCAGCTGCAGGGCAGGCCGGGTGGAGCAGCAGCCACCAGGAGA	60	
Qy	648	CCGGAAACA GTGTGGCCGGCGATCACTCTTGAGCTTGGAGCCAGGAGCCATGTCCTGCCG	707	
Db	61	CCGGAAACA GTGTGGCCGGCGATCACTCTTGAGCTTGGAGCCAGGAGCCATGTCCTGCCG	120	
Qy	708	AGAGATGGCTCAGAGACC CGCGTGATTCGGATGCCCTTCGAAACAGAGCCCTCCCGCAAG	767	
Db	121	AGAGATGGCTCAGAGACC CGCGTGATTCGGATGCCCTTCGAAACAGAGCCCTCCCGCAAG	180	

QY	768	CACGACACAGACAAAGGAGGCGCTCGCTTCCATCTCTAGACGAAAGTACGGCTACTA	827
Db	181	CACGACACAGACAAAGGAGGCGCTCGCTTCCATCTCTAGACGAAAGTACGGCTACTA	240
QY	828	TCACTGCAGAGACTGCAGAAATCCGGGTGAGAGCGCCTATGTGTGTGTGCAGGGCAC	887
Db	241	TCACTGCAGAGACTGCAGAAATCCGGGTGAGAGCGCCTATGTGTGTGTGCAGGGCAC	300
QY	888	CAGTAAGGTGTACTTCAACAGTCTCCGAGTGTGAGAAATCCATCAACCTTACA	947
Db	301	CAGTAAGGTGT-TACTTCAACAGTCTCCGAGTGTGAGAAATCCATCAACCTTACA	359
QY	948	GAGTGAAGACATCACCTGTCAAAAGTTGTAAGAAAGACTAGATGTGCTCCAGTCAGAT	1007
Db	360	GAGTGAAGACATCACCTGTCAAAAGTTGTAAGAAAGACTAGATGTGCTCCAGTCAGAC	419
QY	1008	TTGGCCACGGGAGACCCCTAAAGCCCCCATGGGACAGCTGTGTGGAGATGCAAGGACA	1067
Db	420	TTGGCCACGGGAGACCCCTAAAGCCCCCATGGGACAGCTGTGTGGAGATGCAAGGACA	479
QY	1068	AACGCCATGTCTCTGGACAGGACCTTCAGCTTCAAAATACATCATTTAGTGAAGTGCAGAA	1127
Db	480	AACGCCATGTCTCTGGACAGGACCTTCAGCTTCAAAATACATCATTTAGTGAAGTGCAGAA	539
QY	1128	CGTTTCTGTCTAGATGGGGCTAAATGGAATGGAACAAGTGAAGCTTTTCCCTTTCACTCT	1187
Db	540	CGTTTCTGTCTAGATGGGGCTAAATGGAATGGAACAAGTGAAGCTTTTCCCTTTCACTCT	599
QY	1188	TCCCTTTCCAAATTCCTTCATGACAGACAGTGTACTGTGATATAAGCCTGTGATATAAA	1247
Db	600	TCCCTTTCCAAATTCCTTCATGACAGACAGTGT-TACTGTGATATAAAGCCTGTGATATAAA	658
QY	1248	GGTATTGC	1255
Db	659	GGTATTGC	666

RESULT 3

LOCUS	BC071693	615 bp	mRNA	linear	EST 18-DEC-2003
DEFINITION	H102803.3 N1A Mouse 15K cDNA Clone Set Mus musculus cDNA clone				
ACCESSION	H102803.3				
VERSION	BC071693.2				
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				

ORGANISM

REFERENCE
AUTHORS
Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.U., Wang, X.,
1 (bases 1 to 615)
Bukaryova; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

AUTHORS

TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

TITLE

COMMENT	embryo using a 15,000 mouse developmental CDNA microarray Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
JOURNAL MEDLINE	20381348
PUBLISHED	10922068
ON JAN 26, 2001	this sequence version replaced g1:12554262

COMMENT

Other_EStrs: H3102B03-5
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cna@nigam.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lsun.gsc.nia.nih.gov/CDN/15k.html> for details.

FEATURES

source	1. .615
--------	---------

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="nblast:H3102B03-3"
/db_xref="taxon:10090"
/clone="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI. This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

```

ORIGIN

```

Query Match      46.3%; Score 591.4; DB 10; Length 615;
Best Local Similarity 99.5%; Pred. No. 2.8e-79;
Matches 614; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 643 GAGGACCGGACAGTGTGGCGCGGATGCGTGTGAGCTGGGAGAGGAGCCATGTCCT 702
DB 615 GAGGACCGGACAGTGTGGCGCGGATGCGTGTGAGCTGGGAGAGGAGCCATGTCCT 556
QY 703 GCGGACAGATGGCTCAGGACCCCGGTGATTTGGATGTCCTCGAGACAGGCTCCCGG 762
DB 555 GCGGACAGATGGCTCAGGACCCCGGTGATTTGGATGTCCTCGAGACAGGCTCCCGG 496
QY 763 CAAAGGACGAGAGAGAGAGAGGCGCTGCTCCAGTTCTTAGACAGAGATGAGGC 822
DB 495 CAAAGGACGAGAGAGAGAGAGGCGCTGCTCCAGTTCTTAGACAGAGATGAGGC 436
QY 823 TACTATCACTGCAAGAGATGCAAAATCCGGTGGAGAGCGCCTATGTGTGTGTGAG 882
DB 435 TACTATCACTGCAAGAGATGCAAAATCCGGTGGAGAGCGCCTATGTGTGTGTGAG 376
QY 883 GGCACACAGTAAGTGTACTTCAACAGTTTCCGAGATGTGTGAGAAATCTTCAACC 942
DB 375 GGCACACAGTAAGTGTACTTCAACAGTTTCCGAGATGTGTGAGAAATCTTCAACC 317
QY 943 TTACAGAGTGGAGAGATCACTGTCAAAAGTTTAAAGAACTAGATGTGCTGCCAGT 1002
DB 316 TTACAGAGTGGAGAGATCACTGTCAAAAGTTTAAAGAACTAGATGTGCTGCCAGT 257
QY 1003 CAGATTTGCCGACGTTGAGACCTTAAAGCCGCCATGGCAAGACTTGTGGAGATGCA 1062
DB 256 CAGATTTGCCGACGTTGAGACCTTAAAGCCGCCATGGCAAGACTTGTGGAGATGCA 197
QY 1063 GGCACCAAGCCTGTCTCTGAGACAGACCTTCAAGTTCAATATCATCTTATGAGAGTC 1122
DB 196 GGCACCAAGCCTGTCTCTGAGACAGACCTTCAAGTTCAATATCATCTTATGAGAGTC 137
QY 1123 GAAAACGTTTCTGCTAGATGGGCTTAAAGAAATGAGCAAGTGTGCTTCTCCCTCTTCA 1182
DB 136 GAAAACGTTTCTGCTAGATGGGCTTAAAGAAATGAGCAAGTGTGCTTCTCCCTCTTCA 77
QY 1183 CCTCTTCCCTTCAAAATCTTCAATGACAGACAGTGTACTGATATTAAGCCTGTGAA 1242

```

```

DB 76 CCTCTTCCCTTCAAAATCTTCAATGACAGACAGTGTACTGATATTAAGCCTGTGAA 18
QY 1243 TAAAGTATTCGAAC 1259
DB 17 TAAAGTATTCGAAC 1

```

```

RESULT 4
LOCUS BC084538 595 bp mRNA linear EST 18-DEC-2003
DEFINITION H3102B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION BC084538
VERSION BC084538.2 GI:40072038
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

```

REFERENCE Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
1 (bases 1 to 595)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
Grabovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagareja,R., Doi,H.,
Wood,W.H., III, Becker,K.G., and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
MEDLINE 10922068
PUBMED On Jan 26, 2001 this sequence version replaced gi:12567102.
COMMENT Other ESTs: H3102B03-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
Plate: H3102 row: B column: 03
Seq primer: -21M13 Reverse
High quality sequence stop: 595
POLYA=No.

```

FEATURES

```

source
1..595
Location/Qualifiers
1..595
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="nblast:H3102B03-5"
/db_xref="taxon:10090"
/clone="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI. This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on

```

ORIGIN

the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

Query Match 45.6%; Score 582.4; DB 12; Length 595;
 Best Local Similarity 99.7%; Pred. No. 6.4e-78;
 Matches 594; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 511 CCGGCAATCTCGGGAGACCCGGGAAACCGGACCCGAGAGAGTGGCCCGGAGAAAGCGGTG 570
 DB 1 CCGGCAATCTCTCGGGAGACCCGGGAAACCGGACCCGAGAGAGTGGCCCGGAGAAAGCGGTG 60
 QY 571 CCCGAGCCGCGAAGCGAGAGAGGCGATGTTCAAGCTGACAGGACCGGGTGGAGCA 630
 DB 61 CCCGAGCCGCGAAGCGAGAGAGGCGATGTTCAAGCTGACAGGACCGGGTGGAGCA 120
 QY 631 CAGGCAACCGCGAGAGACCGGAAACAGTGTGGCGCGATGCAAGTCTAGCTCTGAGCGAG 690
 DB 121 CAGGCAACCGCGAGAGACCGGAAACAGTGTGGCGCGATGCAAGTCTAGCTCTGAGCGAG 180
 QY 691 GAGGCAATCTCTCGGGAGAGATGGCTCAAGACCCCGGTATTCGATGCCCCCTCGAGAC 750
 DB 181 GAGGCAATCTCTCGGGAGAGATGGCTCAAGACCCCGGTATTCGATGCCCCCTCGAGAC 240
 QY 751 CAGGCTCTCCCGCAAGCAGAGACAGAGACAGAGAGCGCTGCGTTTCCAGTCTTAGAG 810
 DB 241 CAGGCTCTCCCGCAAGCAGAGACAGAGACAGAGAGCGCTGCGTTTCCAGTCTTAGAG 300
 QY 811 CAGAGATGAGGCTCTATCTACTGCAAGACTGCAAAATCCGCTGGAGAGCGCTTATGTG 870
 DB 301 CAGAGATGAGGCTCTATCTACTGCAAGACTGCAAAATCCGCTGGAGAGCGCTTATGTG 360
 QY 871 TGGGTGTGAGAGGCGACCAAGTGTACTTCTTCAAAACAGTGTGCGAGATGTGAGAA 930
 DB 361 TGGGTGTGAGAGGCGACCAAGTGTACTTCTTCAAAACAGTGTGCGAGATGTGAGAA 419
 QY 931 ATCTCAACCCCTTACAGAGTGAAGACATCACTGTCAAAAGTTGTAAAGAACTAGATG 990
 DB 420 ATCTCAACCCCTTACAGAGTGAAGACATCACTGTCAAAAGTTGTAAAGAACTAGATG 479
 QY 991 TGCTTGGCCAGTCAAGATTTGGCCACGTGAGACCTTAAACGCCCATCCGCAAGACTTGTG 1050
 DB 480 TGCTTGGCCAGTCAAGATTTGGCCACGTGAGACCTTAAACGCCCATCCGCAAGACTTGTG 539
 QY 1051 TGGGAGATGCAAGAGCAAAAGCCTGTCTCTGCAAGACAGCACTTCAAGTCAAAATACA 1106
 DB 540 TGGGAGATGCAAGAGCAAAAGCCTGTCTCTGCAAGACAGCACTTCAAGTCAAAATACA 595

RESULT 5 594 bp mRNA linear EST 04-DEC-2000
 BF471866/c
 LOCUS
 DEFINITION
 UT-M-BH3-awu-b-08-0-UI.r1 NIH BMAP M S4 Mus musculus cDNA clone
 UT-M-BH3-awu-b-08-0-UI 5', mRNA sequence.

ACCESSION
 BF471866.1 GI:11541049
 VERSION
 BF471866.1
 KEYWORDS
 EST.
 Mus musculus (house mouse)
 ORGANSIM
 Mus musculus

REFERENCE
 AUTHORS
 TITLE
 Bernaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890

Email: MEST@mail.nih.gov
 CDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..594
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-awu-b-08-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M S4"
 /note="Vector: pT7D-Pac (Pharmacia) with a modified
 polylinker Site 1: Not I; Site 2: Eco RI; The
 NIH BMAP M S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH BMAP M S4,
 NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,
 NIH BMAP M S2, NIH BMAP M S1. The subtracted library
 (NIH BMAP M S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2 and
 NIH BMAP M S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH BMAP M S4 library. This procedure has been previously
 described (Bernaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)"

ORIGIN

Query Match 44.7%; Score 570.4; DB 10; Length 594;
 Best Local Similarity 99.5%; Pred. No. 4.1e-76;
 Matches 593; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 555 CGCAGAGAAAGCGGTCCCGCAGCGAGAGAGAGGCGATGTTCAAGCTGCAAGGCA 614
 DB 594 CGCAGAGAAAGCGGTCCCGCAGCGAGAGAGAGGCGATGTTCAAGCTGCAAGGCA 535
 QY 615 GCGCGGTGGAGAGACAGCCACCAACCGAGAGACCGGAAACAGTGTGGCGCGATGCACTC 674
 DB 534 GCGCGGTGGAGAGAGACAGCCACCAACCGAGAGACCGGAAACAGTGTGGCGCGATGCACTC 475
 QY 675 TGAGCTTGGAGAGAGAGACATGCTGCGCGAGAGATGCTCAGAGACCCGGTATTC 734
 DB 474 TGAGCTTGGAGAGAGAGACATGCTGCGCGAGAGATGCTCAGAGACCCGGTATTC 415
 QY 735 GGAATGCCCTTGAAGACCAAGGCTCCCGCAAAAGCAGAGAGACCAAGAGCGCTTCG 794
 DB 414 GGAATGCCCTTGAAGACCAAGGCTCCCGCAAAAGCAGAGAGACCAAGAGCGCTTCG 355
 QY 795 TTTCAGTCTTGAAGAGAAAGTACGCTACTATCTGCAAGAGCTGCAAAATCCGGTG 854
 DB 354 TTTCAGTCTTGAAGAGAAAGTACGCTACTATCTGCAAGAGCTGCAAAATCCGGTG 295
 QY 855 GGAAGAGCGCTATGTGTGTGTGCAAGGAGCAGCAAGTAAAGTGTACTTCAAAAGTTCT 914
 DB 294 GGAAGAGCGCTATGTGTGTGTGCAAGGAGCAGCAAGTAAAGTGTACTTCAAAAGTTCT 236

OY		915	GCCGAGTGTGGCAGAAATTCCTCAACACCCTTAACAAGTGAGGAGACATCAACTGTCAAAGTT	974
Dd		225	GCCGAGTGTGGCAGAAATTCCTCAACACCCTTAACAAGTGAGGAGACATCAACTGTCAAAGTT	176
OY		975	GTAAAAGAATCTAGATGTGTGCCTGCCAGTCAGATTTCGCCACGTGGACCCCTAAGCGCCCC	1034
Dd		175	GTAAAAGAATCTAGATGTGTGCCTGCCAGTCAGATTTCGCCACGTGGACCCCTAAGCGCCCC	116
OY		1035	ATCGCAAGACTTGTGTGGAGAGATGCAGAACAAAGCGCTGTCCCGAGACGACCTTCA	1098
Dd		115	ATCGCAAGACTTGTGTGGAGAGATGCAGAACAAAGCGCTGTCCCGAGACGACCTTCA	56
OY		1095	GCTTCAAAATACATCATTTTAGTGAGAGTCGAAAAACGTTTTCTGTAGATGGGGCTAAT	1150
Dd		55	GCTTCAAAATACATCATTTTAGTGAGAGTCGAAAAAC-TTCTGTAGATGGGGCTAAT	1
RESULT 6				
BE704019				
LOCUS				
DEFINITION	BE704019 RIKEN full-length enriched, in vitro fertilized eggs Mus musculus cDNA clone 742045B08 3', mRNA sequence.	525 bp	mRNA	linear EST 11-OCT-2001
ACCESSION	BE704019			
VERSION	BE704019.1			
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	(bases 1 to 525)			
AUTHORS	Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoaka,T., Hirozane,T., Imochi,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Komno,H., Koude,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Saeki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Soabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Taku-Akehira,S., Tanaka,T., Tomaru,A., Toya,T., Watabiki,A., Yasunishi,A., Yumatsuka,M., and Hayashizaki,Y.			
TITLE	RIKEN Encyclopedia of Mouse Full-Length cDNAs (Akimura,T., et al. 2001)			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yoshinide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Tsukuba-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp/ url:http://genome-gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itchih.M., Komno,H., Okazaki,Y., Yamatsuz.M., and Hayashizaki.Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunawa,S., Kawai,J., Okazaki,Y., Yamatsuz.M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Please visit our web site (http://genome-gsc.riken.go.jp/) for further details. e mouse tissues. Location/Qualifiers			
FEATURES				

```

source
1..525
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420459B08"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'] GAGGAGAGAGATCCCAAGACTCTTTTTTTTTTTTTNV 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adaptor of sequence [5'] GAGGAGAGATTCGTGGATTATTAATTAATATCCCCCCCCCCC 3'}. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid pUC(+), after bulk excision from Lambda phage I. Cloning site, 5' end: SalI, 3' end: BamHI"
```

LOCUS AI854700 499 bp mRNA linear EST 15-JUL-1998
DEFINITION U1-M-BHO-akc-d-12-0-U1 S1 NIH BMP M1 Mus musculus cDNA clone
U1-M-BHO-akc-d-12-0-U1 3', mRNA sequence.
ACCESSION AI854700
VERSION AI854700.1 GI:5498606
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 499)
AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (5), 791-806 (1996)
MEDLINE 97044477
PUBMED 8869548
COMMENT Contact: Chin, H

Email: MSB@mail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized basal ganglia library cDNA library Preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
SOLVA=Yes.

FEATURES	Location/Qualifiers
source	1. .499

```

1. .499
/organism="Mus musculus"
/mol_type="mRNA"
/stmol="C57BL/6J"
/db_xref="taxon:10090"
/clone="UT-M-BH0-alc-d-12-0-UI"
/age_stage="2-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_id="N1H_BMAP_M_S1"
/notes="Vector: pTV121-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; site 2: Eco RI. The
N1H BMP M_S1 library is a subtracted library derived from
a mixture of normalized libraries from ten regions of the
mouse brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus). The driver used for
subtraction consisted of a pool of 20 000 cDNA clones
obtained from non-normalized and normalized libraries of
these ten regions of the mouse brain.
TAG_1ISSUE=basal ganglia
TAG_LIB=N1H_BMAP_M_S1
TAG_SEQ=TGATAC"

```

Query Match	37.1%; Score 473.6; DB 9;	Length 499;
Best Local Similarity	99.2%; Pred. No. 1.3e-61;	
Matches 497; Conservative 0;	Mismatches 2;	Gaps 2;
Cy	772 GAGCAGAACAAGAGCGCCCTGTTCAGTTCATTAGACGAATAAGTAGCGACTATCAC	831
Dd	499 GAGCAGCAAGAGAGCGCCCTGTTCAGTTCATTAGACGAATAAGTAGCGACTATCAC	440
Oy	832 TGCAAGAGCTCAAAAATCCGGTGGGAGAGCGCCCTATGTGTGTGTGCAGGCCACCAGT	891
Dd	439 TGCAGAGCTGTAAAAATCCGGTGGGAGAGCGCCCTATGTGTGTGTGTGCAGGCCACCAGT	380

Qy	892	AAGGTGTTACTCTTCAAAACAGTTCTCCGAGTGTGAGAAATCCTACAAACCTTACAGAGT	951
Dp	379	AAGGTG-TACTTCCAAACAGTTCTCCGAGTGTGAGAAATCCTACAAACCTTACAGAGT	321
Qy	952	GGAGGACATCACTGTCAAAGTTGTAAAGAATCTAGATGTGCTGCCAGTCAGATTTG	1011
Dp	320	GGAGGACATCACTGTCAAAGTTGTAAATACTAGATGTGCTGCCAGTCAGACTTG	261
Qy	1012	CCAGGTGAGCCCTTAAAGCCCCCATTCGGCAAGATTTGTGTGGAGATGCAAGGACAAACG	1071
Dp	260	CCAGGTGAGCCCTTAAAGCCCCCATTCGGCAAGATTTGTGTGGAGATGCAAGGACAAACG	201
Qy	1072	CCTGTCCGCGACACGACCTTCGACTTCAATACATCATTTAGAGAGTGGAAAACGTT	1131
Dp	200	CCTGTCCGCGACACGACCTTCACTTCAATACATCATTTAGAGAGTGGAAAACGTT	141
Qy	1132	TCTGTAGATGGGGCTAATGGAATGACAAGTAGAGCTTTCCTCTTCACTCTTCC	1191
Dp	140	TCTGTAGATGGGGCTAATGGAATGACAAGTAGAGCTTTCCTCTCTTCACTCTTCC	81
Qy	1192	TTTTCGAATTTCTCATGACAGACAGTGTACTTGGATATTAAGCTGTGATTAAGGTA	1251
Dp	80	TTTTCGAATTTCTCATGACAGACAGTG-TACTTGGATATTAAGCTGTGATTAAGGTA	22
Qy	1252	TTGCAACCAAAAAAAAAAAAAA 1272	
Dp	21	TTGCAACCAAAAAAAAAAAAAA 1	

[illegible]

ORGANISM Mus musculus
 Eukaryotes: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 491)
 REFERENCE Aitamura, T., Aikawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
 AUTHORS

TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshinide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute for Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-8222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Camircoli, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cdp-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagci, K., Fujitake, S., Inoue, K., Togawa, Y., Iawa, M., Ohara, E.,
Watanuki, M., Yoneda, Y., Ishikawa, T., Osawa, K., Tanaka, T.,
Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, T., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format

Best Local Similarity 86.1%; Pred. No. 5.3e-50;
Matches 464; Conservative 0; Mismatches 71; Indels 4; Gaps 3;

```

QY 713 TGGCTCAGAGACCCCGGTGATTCGATGCCCCCTGAGACCGAGCCCTCCCGGCAAGACCG 772
Db 542 TGGCTCAGAGACCCCGGTGATTCGATGCCCCCTGAGACCGAGCCCTCCCGGCAAGACCG 483
QY 773 AGCAGAGACAGAGAGCGCTGCGTTTCCAGTTCTTAGACAGAGATACGG--CTACTATCA 830
Db 482 AGNACAGAGAGCGCGCTGCGTTTCCAGTTCTTAGACAGAGATACGG--CTACTATCA 423
QY 831 CTGAGAGAGATGCAAAATCCGCTGGAGAGAGCGCTATGTG--TGCTGTGTGACAGGACCA 889
Db 422 CTGAGAGAGATGCAAAATCCGCTGGAGAGAGCGCTATGTG--TGCTGTGTGACAGGACCA 363
QY 890 GTAAAGTGTACTTCAACAGATTGTGCGAGTGTGAGAAATCTTACAAACCTTACAGA 949
Db 362 AGTAAGTGTACTTCAACAGATTGTGCGAGTGTGAGAAATCTTACAAACCTTACAGA 303
QY 950 GTGAGAGACATCAGCTGTCGAAAGTTTAAAGAACTGATGTGCTGCCAGTCAATTT 1009
Db 302 GTGAGAGACATCAGCTGTCGAAAGTTTAAAGAACTGATGTGCTGCCAGTCAATTT 243
QY 1010 CGCCACGTGACACCTTAAACGCCCATCGGAGAGATTGTGAGAGATGCAAGACAA 1069
Db 242 CGCCACGTGACACCTTAAACGCCCATCGGAGAGATTGTGAGAGATGCAAGACAA 183
QY 1070 CGCCTGTCTCGGACAGACCTTCAAGTTCAATATCATCTTATGAGAGTGAAG 1129
Db 182 CGCCTGTCTCGGACAGACCTTCAAGTTCAATATCATCTTATGAGAGTGAAG 123
QY 1130 TTCTGTAGATGAGGCTTAAATGAGATGAGAGATTTCCCTCTTCACTCTTC 1189
Db 122 TTCTGTAGATGAGGCTTAAATGAGATGAGAGATTTCCCTCTTCACTCTTC 63
QY 1190 CTTTCCAAATTTCTTCAATGACAGACAGTGTACTTGATTAAGCTGTGAATAAAG 1248
Db 62 CTTTCCAAATTTCTTCAATGACAGACAGTGTACTTGATTAAGCTGTGAATAAAG 5

```

RESULT 13
LOCUS BE946858 450 bp mRNA linear EST 03-OCT-2000
DEFINITION UI-M-BH3-awu-b-08-0-UI.s1 NIH BMAP M S4 Mus musculus cDNA clone
VERSION BE946858
ACCESSION BE946858.1 GI:10524617
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 450)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL MEDLINE
PUBMED 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mst@mail.nih.gov

Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements

Seq primer: M13 Forward
POLYA=No.

FEATURES
source location/Qualifiers

```

1..450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-awu-b-08-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_id="NIH_BMAP_M_S4"
/note="Vector: pT73D-Pac (pharmacia) with a modified
NIH_BMAP_M_S4 library is a subtracted library of a series
of polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP_M_S4 library is a subtracted library of a series
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stem, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S1,
NIH_BMAP_M_S2, NIH_BMAP_M_S3, NIH_BMAP_M_S4,
NIH_BMAP_M_S5, NIH_BMAP_M_S6, NIH_BMAP_M_S7,
NIH_BMAP_M_S8, NIH_BMAP_M_S9, NIH_BMAP_M_S10,
NIH_BMAP_M_S11, NIH_BMAP_M_S12, NIH_BMAP_M_S13,
NIH_BMAP_M_S14. The subtracted library
cDNA inserts from NIH_BMAP_M_S3, NIH_BMAP_M_S4, and
NIH_BMAP_M_S5 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3, NIH_BMAP_M_S4, and NIH_BMAP_M_S5.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M_S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_SEQ=None found"

```

ORIGIN

Query Match 30.8%; Score 393; DB 10; Length 450;
Best Local Similarity 99.7%; Pred. No. 1.9e-49;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 3 GGGGGGGAGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 62
Db 57 GGGGGGGAGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 116
QY 63 TCCTTATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 122
Db 117 TCCTTATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 176
QY 123 ACCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 182
Db 177 ACCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 236
QY 183 CAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 242
Db 237 CAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 296
QY 243 CAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 302
Db 297 CAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 356
QY 303 CAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
Db 357 CAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 416
QY 363 CTTGTCAACCCCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
Db 417 CTTGTCAACCCCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 450

```

RESULT 14
BB704449
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BB704449 419 bp mRNA linear EST 11-OCT-2001
BB704449 RIKEN full-length enriched, in vitro fertilized eggs Mus.
musculus cDNA clone 7420464A04 3', mRNA sequence.
BB704449
EST.
BB704449.1 GI:16053284

Mus musculus (house mouse)
Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 419)

REFERENCE
AUTHORS

1 Akimura, T., Aikawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,
Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Komuro, H., Kouda, M.,
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nunasaki, R.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Tanaka, T., Tomaru, A., Toya, T., Watanishi, A., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
2001)

TITLE

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saito-cho, Tsukuba-shi, Ibaraki, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
url: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

Wagui, K., Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.,
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Komuro, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES
source

Location/Qualifiers

1. 419
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420464A04"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clone_1fb="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: Salt, Site 2: BamHI. cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken

ORIGIN

Query Match 30.7%; Score 392.2; DB 10; Length 419;
Best Local Similarity 98.8%; Pred. No. 2,5e-49;
Matches 416; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 834 CAAGACTGCAAAATCCGGTGGAGAGAGCCCTATGCTGGTGTGTGAGGGACCCAGTA 893
DB 1 CAAGACTGCAAAATCCGGTGGAGAGAGCCCTATGCTGGTGTGTGAGGGACCCAGTA 60

QY 894 GGTGTACTTCAACACAGTTCTGCGAGTGTGTGAGAAATCTTCAACCCCTTACAGAGTGG 953
DB 61 GGT-TTACTTCAACACAGTTCTGCGAGTGTGTGAGAAATCTTCAACCCCTTACAGAGTGG 119

QY 954 AGGACATCAGCTGTGCAAAAGTTTAAAGAACTGATGTGCTGCGCCAGTATGTCGCC 1013
DB 120 AGGACATCAGCTGTGCAAAAGTTTAAAGAACTGATGTGCTGCGCCAGTATGTCGCC 179

QY 1014 ACGTGACCTTAAAGCCGCCCATCGCAAGCTTGTGTGAGAGATGCAAGACAAAGCCG 1073
DB 180 ACGTGACCTTAAAGCCGCCCATCGCAAGCTTGTGTGAGAGATGCAAGACAAAGCCG 239

QY 1074 TGTCTGCGACAGCAGCTTCAAGCTTCAATATCATTTATGTAGAGTCAAAACGTTTC 1133
DB 240 TGTCTGCGACAGCAGCTTCAAGCTTCAATATCATTTATGTAGAGTCAAAACGTTTC 299

QY 1134 TGTCAATGAGGGGCTAATGAAATGAAAGCAAGAGCTTTCCTCCCTTCACTCTTCCTT 1193
DB 300 TGTCAATGAGGGGCTAATGAAATGAAAGCAAGAGCTTTCCTCCCTTCACTCTTCCTT 359

QY 1194 TCCAAATCTTCAATGACAGCAGTGTACTTGGATATTAAGCCTGTGAATTAAGGTATT 1253
DB 360 TCCAAATCTTCAATGACAGCAGTGTACTTGGATATTAAGCCTGTGAATTAAGGTATT 418

QY 1254 G 1254
DB 419 G 419

RESULT 15
BB703869
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BB703869 521 bp mRNA linear EST 11-OCT-2001
BB703869 RIKEN full-length enriched, in vitro fertilized eggs Mus.
musculus cDNA clone 7420457C21 3', mRNA sequence.
BB703869
EST.
BB703869.1 GI:16052704

Mus musculus (house mouse)
Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 521)

REFERENCE
AUTHORS

1 Akimura, T., Aikawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,
Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Komuro, H., Kouda, M.,
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nunasaki, R.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Tanaka, T., Tomaru, A., Toya, T., Watanishi, A., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.

JOURNAL	the oocyte-to-embryo transition
MEDLINE	Nat. Genet. 33 (2), 187-191 (2003)
PUBMED	22447938
AUTHORS	15359046
TITLE	2 (Pages 1 to 1260)
JOURNAL	Wu X., Wang, P. and Matczuk, M.M.
FEATURES	Direct Submission
SOURCE	Submitted (04-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA Location/Qualifiers 1. 1260

3' UTR	1114, 1260
ORIGIN	/gene="Zar1"

423 CCGTCTCGTCCGTGACCTTCTGTGGCCCTCCTCTCTACTGAGGTTGCGGAGGCAG 482

[illegible]

MEDLINE 22811438
 PUBMED 12773403
 REFERENCE 2 (pages 1 to 1280)
 AUTHORS Wu, X. and Matzruk, M.M.
 TITLE Direct Submission
 JOURNAL Submitted (24 APR 2003) Pathology, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 FEATURES Location/Qualifiers
 source 1. .1280

ORIGIN

Query Match	75.3%;	Score 962.2;	DB 10;	Length 1280;
Best Local Similarity	87.7%;	Pred. No. 1.9e-189;		
Matches 110;	Conservative	0;	Mismatches 143;	Indels 12;
				Gaps 5;

QY	1	GGGCGGGGAGAGGGCGGGGAGACGACCCATGTTTCCCGGCGAGACAGTTTCAACCCCTGCCCGCA	62
Db	3	GCGCGCGAGGCGCGGGAGCGACCCATGTTTCCCGGCGAGACGCCCCAACCCATGCCCGCA	62
QY	63	TCCTTATCCGC--AGGCCACCAAAAGCCGGGATGGCTGAGGTTTCGAGGCCAGGGCTG	119
Db	63	TCCTTACCGGCCACGGCAGCCAAAGCGGGGATGGCTGAGGTTTGAAGCAGGGGCTG	122
QY	120	CGAACCCGGGCCCCCTCTTCTTCTTCCCGGCTACAGACAGTCAATGGCCGCGAGATACG	179
Db	123	CAGGCCGGAGCCCCCTCTTCTTCTTCCCGGCTACAGACAGTCAATGGCCGCGAGATAC	182
QY	180	CGACAGCCACCAAGCGGGGACAGCATATGAGCCCTGTCGCGAATGGATCCCGGATCGGT	239
Db	183	TGACACTATACAGCGAGGCGGCACTATAGGCTTGTGTGCGGAATGGGTCGCCGCGCGGT	242
QY	240	CAGCAGCCTGTAGCGCTGCGGTGCGAGTGAAACCCGCGCCGACGCTTCGGTGACATGTTT	299
Db	243	CACACACCGCGAGAGCTGCGGTGCAAGTAGAACCCGCGCGCATGCTCGGTGACATGTTT	302
QY	300	ACTCGGCGCGCGACAGCTGACGCTGCAAGGATGCCAGGCCGACCGCCGATGCGG	359
Db	303	GCTCGGGGCGCGACACTGCAAGCTTGAGACGAGGCCGACAGCCCGACGCGCGCTGG	362
QY	360	TTCCTGTCAAACCCCGTGGGACAGCGCGCGCGCGGAGATCCCGCGATCTTGGACGACCGT	419
Db	363	TTCCTGCAAAACCCCGAGCGCGCGCGCGGAGACCCCGCGATCTTGGCGCACCGT	422
QY	420	AGCGCCGTTCTGTGCTCGTGACCTTGTGTGAGCTTCCCTCCCTCACTGGAGAGTTGCCGAGG	479
Db	423	CGGCTGTACTGCGCCCGTGAACCTTGGTGGCTCTCTCTCTGCTGAGGTTGGGGGGA	482
QY	480	CAGGCAAGACCAACGAGAGGAGAGAGGAGGCCCGCATCTTGGGAGACCCGAGAACCGAG	539
Db	483	CAGGCAAGACCGCCACGAGAGGAGAGGAGGACCGGCACTCACGGGAGCCCGGAGACCGGA	542
QY	540	GCGGAGAGAGTGGCCGCGAGGAAAGCGGTCCCCAGCGCGGAAAGCGAGAGGCGCATGT	599

Db	543	GCCGGAGAGAGTGGCAGTGAATGAAAGACAGTCCCCACGCCGACAGAGCGAGGCGACGT	602
Qy	600	TCAGGCTTCAGAGGCAGAGCGCGGATGGAGACAGCACAGCCACACCAGAGGACCCGMAACAGTGT	659
Db	603	CCAGGCTGMAAGGAGAGATG---GGACAGAACACACCCACCGGGAGAGACCCGGACACAGTGT	659
Qy	660	GGCGGCGATGACGTCTGAGCCTTGGGAGCGAGAGACCAATGTCTCTCCGCGAGATAGACTCA	719
Db	660	GGCGGCGATGACGTCTGAGCCTTGGGAGCGAGAGACCACTCTCTCTCTGCGAGATAGCTCA	719
Qy	720	GGAACCCGGGTATTTGGATGCGCTTCGAGACAGGCTTCCCGCGAAGACAGAGACAGA	779
Db	720	GGACCCCGATGACGTGGCTTCTTGAAGACCGGAGCTCCCGACAGAGCACTGAGCAGA	779
Qy	780	CAAGAGAGGCTGGGTTTCCAGTCTTAAAGCAAAAGTACGGCTACTTCACTGCAAGGA	839
Db	780	CAAGAGAGGCTGGGTTTCCAGTCTTAAAGCAAAAGTACGGCTACTTCACTGCAAGGA	839
Qy	840	CTGCAAAATCCGATGGAGAGAGCGCCCTATGTGTGTGTGTGTGAGAGGACACAGTAAAGTGT	899
Db	840	CTGCAAAATCCGATGGAGAGAGCGCCCTATGTGTGTGTGTGTGAGAGGACACAGTAAAGTGT	899
Qy	900	ACTTCAAAACAGTTTCTGCCGAGTGTGTGAGAAATCCTTACAACCTTTACAGATGAGAGCA	959
Db	899	ACTTCAAAACAGTTTCTGCCGAGTGTGTGAGAAAGTCTTACAACCCATACCGATGAGAGATA	958
Qy	960	TCACCTGTCAAAGTGTGAAAGAACTGATGTGCTCCGCCAGTACAGATTTTGGCACAGTG	1019
Db	959	TCACCTGTCAAAGTGTGAAAGAACTGATGTGCTCCGCCAGTACAGATTTTGGCACAGTG	1018
Qy	1020	ACCTTAAACGCCCCCATATGCGCAGACCTTGTGTGGAGATGACAGACAAACGCTGTCTCT	1079
Db	1019	ACCTTAAACGCCCCCATATGCTTCAACACTTGTGTGGAGATGACAGACAAACGCTGTCTCT	1078
Qy	1080	GCGACAGACCTTACGCTTCAAAATACATCTTGTGAGAGTGCAGAAACGTTTCTGTCTA	1138
Db	1079	GTCACAGACCTTACGCTTCAAAATACATCTTGTGAGAGTGCAGAAATGTTTCTGTCTA	1138
Qy	1139	---GATGGGGCTAATGAAATGAGACAACTGACCTTCTCCCTCTTCACTTCCCTTC	1199
Db	1139	ATGATGATGGAACAAATGAGACCTTCTCCACAGCGCCCTCTCCCTCTTCACTTCCCTTC	1198
Qy	1196	CAATTTCTTACATGACACACACTGTACTTGTGATTAAGCTCTGAAATTAAGATATTGC	1255
Db	1199	AAATTAATCTTACATGAAAGCAGTG-TACTTTAATATTAAGCTCTGAAATTAAGAGGATTGC	1257
Qy	1256	AAACA 1260	
Db	1258	AAACA 1262	
RESULT 3			
AY193889	3987 bp	DNA	linear ROD 12-MAR-2003
LOCUS			
DEFINITION	Mus musculus zygote arrest 1 (Zarl) gene, complete cds.		
ACCESSION	AY193889		
VERSION	AY193889.2	GI:27808699	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Mu,X., Viveiros,M.M., Eppig,J.J., Bai,Y., Fitzpatrick,S.L. and Mazzuk,M.M.		
TITLE	Zygote arrest 1 (Zarl) is a novel maternal-effect gene critical for the oocyte-to-embryo transition		
JOURNAL	Nat. Genet. 33 (2), 187-191 (2003)		
PMID	22447938		
PMED	12539046		
REFERENCE	2 (bases 1 to 3987)		
AUTHORS	Mu,X. and Mazzuk,M.M.		
TITLE	Direct Submission		

JOURNAL Submitted (09-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 3987)
 AUTHORS Wu, X. and Matzuk, M.M.
 TITLE Direct Submission
 JOURNAL Submitted (21-JAN-2003) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REMARK Sequence update by submitter
 COMMENT On Jan 21, 2003 this sequence version replaced gi:27808693.
 FEATURES
 source
 1. .3987
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129/SvEv"
 /db_xref="taxon:10090"
 1. .3987
 /gene="Zar1"
 /join(1..801,2574..2666,2745..2819,3697..3987)
 /gene="Zar1"
 /product="zygote arrest 1"
 1. .801
 /gene="Zar1"
 /number=1
 /join(28..801,2574..2666,2745..2819,3697..3840)
 /gene="Zar1"
 /note="oocyte-specific"
 /codon_start=1
 /product="zygote arrest 1"
 /protein_id="AA024708.1"
 /db_xref="GI:27808694"
 /translation="MPPSTFHPCHPYPOATKADGWRFGARCPAPSPFLPGYRQ LMAAEVDSHORALMLSRMGPVSRSRDAVOVPRPDASVCGSGRLLDOPAGC RASPDARSGSCOPRGAGAGRS PRSMOTUAPSSYTFGCSGLSEVAGRPTPKGSE SPASSGTREPERREYVAAKRAVPOPRSESDVQAAGACMEQPPREDNYSVAADKQSE GSEPCPAEWAADPEDSDAPRDQASPOSTEDKRLRFQLEQKRYGHCKDKXIRM ESAVWCVQGISKRYFKQPCRVCEKSYNPRVEDITCQSCKTRCACPVRLRHVDPKR PHRODLGRCKRKLSCDSTSFXYI1"
 2574..2666
 /gene="Zar1"
 /number=2
 2745..2819
 /gene="Zar1"
 /number=3
 3697..3987
 /gene="Zar1"
 /number=4
 ORIGIN
 Query Match 62.6%; Score 799.4; DB 10; Length 3987;
 Best Local Similarity 99.9%; Pred. No. 1,2e-155;
 Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 3 GGGGGGCGAGGCGGCGGAGCGACCCATGTTCCCGGCGAGCAGCTTCCACCCCTGCGCGCA 62
 3 GGGGGGCGAGGCGGCGGAGCGACCCATGTTCCCGGCGAGCAGCTTCCACCCCTGCGCGCA 62
 63 TCTTATTCGCGAGGCGACCAAAACCGGGATGCTGGAAGTTTCGAGCCAGGCGCTGCGG 122
 63 TCTTATTCGCGAGGCGACCAAAACCGGGATGCTGGAAGTTTCGAGCCAGGCGCTGCGG 122
 63 TCTTATTCGCGAGGCGACCAAAACCGGGATGCTGGAAGTTTCGAGCCAGGCGCTGCGG 122
 123 ACCCGGCGCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTGA 182
 123 ACCCGGCGCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTGA 182
 123 ACCCGGCGCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTGA 182
 183 CAGCCACGAGGCGGCGACAGCTCATGCGCTGTGCGGATGGATGCCCGGTGCGTCA 242
 183 CAGCCACGAGGCGGCGACAGCTCATGCGCTGTGCGGATGGATGCCCGGTGCGTCA 242
 183 CAGCCACGAGGCGGCGACAGCTCATGCGCTGTGCGGATGGATGCCCGGTGCGTCA 242
 243 CAGCCGTAAGCTGCGGCTGAGTGAACCGCGCGCGACGCTTGGTGAGTGTCTACT 302
 243 CAGCCGTAAGCTGCGGCTGAGTGAACCGCGCGCGACGCTTGGTGAGTGTCTACT 302
 303 CCGGCGCGGCGACGCTGACGCTGCGAGGGTCCGAGCCAGCCCGCATCGGCTTC 362

Db 303 CCGGCGCGGCGACGCTGAGCTGAGAGGTGCCAGCGCCGCGACGCCGATCGGGTTC 362
 Qy 363 CTGTCAACCCCGTGGGCGGAGCGCGCGGCGGAGATCCCGCGATCTCGGAGCCGATG 422
 Db 363 CTGTCAACCCCGTGGGCGGAGCGCGCGGCGGAGATCCCGCGATCTCGGAGCCGATG 422
 Qy 423 CCCGTTCTGCTCGTGAACCTTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 Db 423 CCCGTTCTGCTCGTGAACCTTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 Qy 483 GCAACACCCACGAGGAGAGAGGAGAGCCCGGATCTCTCGGAGACCCCGGAGACCC 542
 Db 483 GCAACACCCACGAGGAGAGAGGAGAGCCCGGATCTCTCGGAGACCCCGGAGACCC 542
 Qy 543 GAGAGAGTGGCGCGCGAGAAACCGGCTCCCGGCGCGAGAGAGAGAGAGAGATGTTCA 602
 Db 543 GAGAGAGTGGCGCGCGAGAAACCGGCTCCCGGCGCGAGAGAGAGAGAGAGATGTTCA 602
 Qy 603 GAGTGCAGGAGAGAGCGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
 Db 603 GAGTGCAGGAGAGAGCGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
 Qy 663 GCGCATGCACTGAGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
 Db 663 GCGCATGCACTGAGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
 Qy 723 CCGCGTGTATTCGATGCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
 Db 723 CCGCGTGTATTCGATGCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
 Qy 783 GAGAGCGCTGCTGTTCCAGTT 803
 Db 783 GAGAGCGCTGCTGTTCCAGTT 803
 RESULT 4
 AC122733
 LOCUS
 DEFINITION Mus musculus clone RP24-506B15, WORKING DRAFT SEQUENCE, 21
 AC122733
 UNORDERED PIECES.
 AC122733
 VERSION 1
 HTG: HTG6 PHASE1, HTG6 DRAFT.
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 1 (bases 1 to 147464)
 Birren, B., Nussbaum, C. and Lander, E.
 Mus musculus, clone RP24-506B15
 Unpublished
 2 (bases 1 to 147464)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalil, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choehel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dehrellano, K., Dewar, K., Diaz, J., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kels, C., Labèque, K., Lamazares, R., Lander, T., Lehotzky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menes, J., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retra, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stefanovic, N., Straus, N., Subramanian, A., Talmas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,

TITLE
JOURNAL
REFERENCE
AUTHORS

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 147464)
Biren, B., Nusbaum, C., Lander, E., Abouelella, A., Allen, N., Anderson, S., Arachchi, H. M., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Decker, J., Fero, S., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fano, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hater, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanab, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meidirim, J., Menus, L., Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, V., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (02-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 2, 2003 this sequence version replaced g121026317.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL
COMMENT

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: L17506
Center clone name: 506_B_15
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 140214 bases at least Q40
Consensus quality: 143454 bases at least Q30
Consensus quality: 144691 bases at least Q20
Insert size: 144000; agarose-fp
Insert size: 145464; sum-of-coverage
Quality coverage: 4.6 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-coverage
NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1
654 753: contig of 653 bp in length
754 1934: contig of 1181 bp in length
1935 2034: gap of 100 bp
2035 3196: contig of 2962 bp in length
3197 3206: gap of 100 bp
3207 3298: contig of 890 bp in length
3299 3308: gap of 100 bp
3309 3381: contig of 725 bp in length
3382 3391: gap of 100 bp
3392 3542: contig of 1518 bp in length
3543 3552: gap of 100 bp

FEATURES
source
1..147464
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_lib="RP24-506B15"
/clone_1ib="RP24-506B15"
1..653
/note="assembly_fragment"
754..1934
/note="assembly_fragment"
2035..3196
/note="assembly_fragment"
3207..3298
/note="assembly_fragment"
3309..3381
/note="assembly_fragment"
3392..3542
/note="assembly_fragment"
3553..3815
/note="assembly_fragment"
3826..4158
/note="assembly_fragment"
4169..4321
/note="assembly_fragment"
4332..4693
/note="assembly_fragment"
4704..5140
/note="assembly_fragment"
5141..5144
/note="assembly_fragment"
5145..5635
/note="assembly_fragment"
5636..6191
/note="assembly_fragment"
6192..7053
/note="assembly_fragment"
7054..7663
/note="assembly_fragment"
7664..8617
/note="assembly_fragment"
8618..9417
/note="assembly_fragment"
9418..10852
/note="assembly_fragment"
10853..12379
/note="assembly_fragment"
12380..14691
/note="assembly_fragment"
14692..14706
/note="assembly_fragment"
14707..14746
/note="assembly_fragment"

misc_feature
/note="assembly_fragment"
754..1934
/note="assembly_fragment"
2035..3196
/note="assembly_fragment"
3207..3298
/note="assembly_fragment"
3309..3381
/note="assembly_fragment"
3392..3542
/note="assembly_fragment"
3553..3815
/note="assembly_fragment"
3826..4158
/note="assembly_fragment"
4169..4321
/note="assembly_fragment"
4332..4693
/note="assembly_fragment"
4704..5140
/note="assembly_fragment"
5141..5144
/note="assembly_fragment"
5145..5635
/note="assembly_fragment"
5636..6191
/note="assembly_fragment"
6192..7053
/note="assembly_fragment"
7054..7663
/note="assembly_fragment"
7664..8617
/note="assembly_fragment"
8618..9417
/note="assembly_fragment"
9418..10852
/note="assembly_fragment"
10853..12379
/note="assembly_fragment"

	*	1	3391:	contig of 3391 bp in length
	*	3392	3491:	gap of 100 bp
	*	3492	11221:	contig of 7730 bp in length
	*	11222	11321:	gap of 100 bp
	*	11322	20171:	contig of 8650 bp in length
	*	20172	20271:	gap of 100 bp
	*	20272	45821:	contig of 25550 bp in length
	*	45822	45921:	gap of 100 bp
	*	45922	100777:	contig of 54856 bp in length
	*	100778	100877:	gap of 100 bp
	*	100878	143469:	contig of 42592 bp in length
	*	143470	143569:	gap of 100 bp
	*	143570	168073:	contig of 24504 bp in length.
FEATURES			Location/Qualifiers	
source			1..168073	
			/organism="Mus musculus"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:10090"	
			/chromosome="5"	
			/map="5"	
			/clone="RP24-22B12"	
			/clone_lib="RPC1-24 Male Mouse BAC"	
misc_feature			1..3391	
			/note="assembly_fragment"	
			clone_end:SP6	
			"vector_side:left"	
misc_feature			3492..11221	
			/note="assembly_fragment"	
misc_feature			11322..20171	
			/note="assembly_fragment"	
misc_feature			20272..45821	
			/note="assembly_fragment"	
misc_feature			45922..100777	
			/note="assembly_fragment"	
misc_feature			100878..143469	
			/note="assembly_fragment"	
misc_feature			143570..168073	
			/note="assembly_fragment"	
			clone_end:T7	
			"vector_side:right"	
ORIGIN				
Query Match			54.0%;	Score 689.8; DB 2; Length 168073;
Best Local Similarity			95.0%;	Pred. No. 6,7e-133;
Matches 761; Conservative			0; Mismatches 22; Indels 16; Gaps 4;	
QY	3	GGCGGCGAGGCGCGGGAACGCATCTTCCGGCAGACAGTCCTCACCCCTGCCGCA	62	
Dd	63451	GGCGGGGAGGCGCGGGAACGCATCTTCCGGCAGACAGTCCTCACCCCTGCCGCA	63510	
QY	63	TCTTTATCCGAGGSCACCAAAGCGGGAGATGTGTGAGATTGAGACCAGGGGCTGGCG	122	
Dd	63511	TCTTTATCCGAGGSCACCAAAGCGGGAGATGTGTGAGATTGAGACCAGGGGCTGGCG	63570	
QY	123	ACCGCGCCCCCTCTCTCTCTCCCGGACTACAAGCTATAGCCCGGAGTAGCTGCA	182	
QY	63571	ACCGCGCCCCCTCTCTCTCTCCCGGACTACAAGCTATAGCCCGGAGTAGCTGCA	63630	
Dd	183	CAGGCACAGGAGGSCAAGCTCATATGCGCTGTGTGCGGATGGATGCCCGGTCAGTCA	242	
QY	63631	CAGGCACAGGAGGSCAAGCTCATATGCGCTGTGTGCGGATGGATGCCCGGTCAGTCA	63690	
Dd	243	CAGCGCTGAGCGTGCGGTGCAAGTGAACCCGCGCCGAGATGCTCGGTGAGTGTCACT	302	
QY	63691	CAGCGCTGAGCGTGCGGTGCAAGTGAACCCGCGCCGAGATGCTCGGTGAGTGTCACT	63750	
Dd	303	CGGGCGCGGACGCTGAGCCTTGAGGATGCCAGAGCAGCCCGAGACGCCGATCGGGTTC	362	
QY	63751	CGGGCGCGGACGCTGAGCCTTGAGGATGCCAGAGCAGCCCGAGACGCCGATCGGGTTC	63810	
Dd	363	CTGTCAACCCCCTGTGACAGCCCGCGCGCGGGAGAATCCCGCATCTTGGACAGCGGTAC	422	
QY	63811	CTGTCAACCCCCTGTGACAGCCCGCGCGCGGGAGAATCCCGCATCTTGGACAGCGGTAC	63870	

Db	Accession	Definition	Accession	Keywords	Source	Organism
Oy	423	CCCGTTCTCGTCCGTGACCTTCTGTGGACCTCTCTCTCACTGGAGGTTCCGGAGGAC	482			
Db	63871	CCCGTTCTCGTCCGTGACCTTCTGTGGACCTCTCTCTCACTGGAGGTTCCGGAGGAC	63922			
Oy	483	GCAACACCCACGAGAGGAGAGGAGGAGCCCGGACATCTGTGGAGACCCCGGAGACCCGAGCC	542			
Db	63926	GCAACACCCACGAGAGGAGAGGAGGAGGAGCCCGGACATCTGTGGAGACCCCGGAGACCCGAGCC	63988			
Oy	543	GAGAGAGGTGAGCGCGCGAGGAAGCGGTCCCGACCGCGAGAGGAGGAGGCGCATGTTCA	602			
Db	63986	GAGAGAGGTGAGCGCGCGAGGAAGCGGTCCCGACCGCGAGAGGAGGAGGCGCATGTTCA	64044			
Oy	603	GGCTGACGAGGAGCGCGGTGGAGAGCAGACGCCACCGGAGAGACCGGAGACAGTGTGGC	662			
Db	64046	GGCTGAGAGGAGCGCGCGGTGGAGAGCAGACGCCACCGGAGAGACCGGAGACAGTGTGGT	64100			
Oy	663	GGCGATCAGTCTTGAGCTTGGAGCGAGAGGACCATGTCTTGGCGAGAGATGGCTCAGGA	722			
Db	64106	GGTATATCAGTCTTGAGCTTGGAGCGAGAGGACCATGTCTTGGCGAGAGATGGCTCAGGA	64165			
Oy	723	CCCCGGTATTCGGATGCCCTCGAGACGACGCGCTCCCGGCAAGACCGAGAGAGCA	782			
Db	64166	CCCCGGTATTCGGATGCCCTCGAGACGACGCGCTCCCGGCAAGACCGAGAGAGCA	64211			
Oy	783	GAGAGCGCTTCGCTTCCAGTT	803			
Db	64213	GAGAGCTTCCTCGCTTCCAGTT	64233			
RESULT 7						
AC126519/c						
LOCUS						
AC126519	235390 bp	DNA	linear	HTG 20-NOV-2002		
DEFINITION	Rattus norvegicus clone CH230-159N5, WORKING DRAFT SEQUENCE.					
AC126519	AC126519.3	GI:25138273				
VERSION	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.					
KEYWORDS	Rattus norvegicus (Norway rat)					
SOURCE	Rattus norvegicus					
ORGANISM	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE						
AUTHORS	1 (bases 1 to 235390)					
	Wuzy, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alshrocks, S., Amiri, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benmhad, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, U., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crete, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escoto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., Gabisi, A., Ganra, R., Garcia, A., Garner, T., Garza, M., Gbragoge, E., Geer, J., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hlaun, S., L., Hodgson, A., Hogues, J., Hollins, B., Howells, S., Huylk, S., Hune, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpich, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowicz, C., Kraft, C., L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, D., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, K., Louie, L., Louie, L., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., P., McNeill, T., Z., Meenen, S., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundae, M., Murphy, M., Nair, L.,					


```

OY      660  GGCGGCATGACATCTGAGCTGGAGCGAGAGCCATGTCTCTCCGACAGATGGCTCA 719
DB      26812 GGGCGGCATGACATCTGAGCTGGAGCGAGAGCCATCTCTCTCTCAAGATGGCTCA 26753
OY      720  GGAGCCCGGATGATCGATGCTCCCTCGAGACAGAGCCCTCCCGCAAGACAGAGAGCA 779
DB      26752 GGAGCCCGGATGATCGATGCTCCCTCGAGACAGAGCCCTCCCGCAAGACAGAGAGCA 26693
OY      780  CAAGAGAGCGCTGCTGCTTCCAGTT 803
DB      26692 CAAGAGAGCGCTGCTGCTTCCAGTT 26669

RESULT 8
AC125993/c
LOCUS   237695 bp  DNA  linear  HTG 10-MAY-2003
DEFINITION
Rattus norvegicus clone CH230-74111, *** SEQUENCING IN PROGRESS
AC125993
VERSION AC125993.3 GI:30522839
KEYWORDS HTG, HTGS, PHASE2, HTGS, DRAFT, HTGS_ENRICHED.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
          1 (bases 1 to 237695)
REFERENCE
Munty D. Marie, Metzker M. Lee, Abramson S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anjalaebchi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benham F.,
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cadenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cre A., D'Souza L.,
Devila M. L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denon S., Deramo C., Ding Y., Dinh H., Divya K.,
Diaper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Escotto M., Eugene C., Evans C. A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster T., Foster P.,
Fraser C. M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Georgescu E., Geer K., Gill R., Grady M., Guerra M., Guevara W.,
Gunnarson P., Haaland M., Hamill C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S. L., Hodgson A., Hognes M.,
Hollins B., Howell S., Huylk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
Karpachy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
Kowis C., Kraft C. L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
Lorenzetti W., Louisedge H., Lozano R. J., Lu X., Ma J.,
Maheshwari M., Mahindaratne M., Mahmoud M., Mallory K., Mangum A.,
Mangum B., Mapa P., Martin K., Martin R., Martinez E.,
Mawhney S., McLeod M. P., McNeill T. Z., Meenen E.,
Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
Morgan M., Morris K., Morris S., Munidasa M., Murphy M., Nair L.,
Nankervis C., Neal D., Newton N., Nguyen N., Norris S.,
Nankovskiy O., Okunou G., Olariuunsgoon A., Pal S., Parks K.,
Pasternak S., Paul H., Perez A., Perez L., Plankoch C.,
Plappert F., Polidexter A., Popovic D., Pitman E., Pu L., L.,
Plopper F., Quiroz J., Rachlin E., Reeves K., Regier M. A., Reihl R.,
Riley B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F.,
Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S. J.,
Sanders W., Savary S., Scherer S., Scott G., Shlaman S., Shen H.,
Shetty J., Shvartsberg A., Sisson I., Sitter C. D., Smajic D.,
Sneed A., Sodergren E., Song X. Z., Sorrell R., Soth A.,
Steinle M., Strong R., Sutton A., Svatek A., Taber P., Taylor C.,
Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Uman K.,
Valas R., Vera V., Villaseana D., Waldron L., Walker B., Wang J.,
Wang Q., Wang S., Warren J., Warren R., Wei X., White F.,
Williams G., Willson R., Wleczky R., Woodson H., Worley K.,
Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
AUTHORS
REFERENCE
JOURNAL
COMMENT
Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von
Niederhausern A., Weiss R., Smith D. R., Holt R. A., Smith H. O.,
Weinstock G. and Gibbs R. A.
Direct Submission
2 (bases 1 to 237695)
Unpublished
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 237695)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23096451.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G61Z
Center clone name: CH230-74111
----- Summary Statistics
Assembly program: Atlas 3.0
Consensus quality: 229641 bases at least Q40
Consensus quality: 231444 bases at least Q30
Consensus quality: 232679 bases at least Q20
Estimated insert size: 242786; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 237695: contig of 237695 bp in length.
----- Location/Qualifiers
1. 237695
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-74111"
1. 1905
/notes="wgs_end_extension
clone_end:T7"
complement(5394..6172)
/notes="clone boundary
clone_end:T7
site:Ecort
end_sequence:BH339391"
complement(234275..235086)

```

FEATURES

source

misc_feature

misc_feature

misc_feature

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mhova, T., Mngwa, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
(bases 1 to 212848)

Anderson, S., Barna, N., Baerlein, V., Bloom, T., Boguslavsky, U., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mhova, T., Mngwa, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 28, 2002 this sequence version replaced gi:18450109.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Center code: W1BR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information
Center project name: L21014
Center clone name: 300_K5

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 200069 bases at least Q40
Consensus quality: 206302 bases at least Q30
Consensus quality: 208218 bases at least Q20
Insert size: 209348; sum-of-contigs
Quality coverage: 5.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

FEATURES

* the accession number will be preserved.
1
641: contig of 641 bp in length
742
741: gap of 100 bp
1427
1427: contig of 686 bp in length
1428
1527: gap of 100 bp
1528
3170: contig of 1643 bp in length
3171
3270: gap of 100 bp
3271
4396: contig of 1126 bp in length
4397
4496: gap of 100 bp
4497
5909: contig of 1413 bp in length
5910
6009: gap of 100 bp
6010
7583: contig of 1574 bp in length
7584
7683: gap of 100 bp
7684
8975: contig of 1292 bp in length
8976
9075: gap of 100 bp
9076
10257: contig of 1182 bp in length
10258
10357: gap of 100 bp
10358
11608: contig of 1251 bp in length
11609
11708: gap of 100 bp
11709
13289: contig of 1581 bp in length
13290
13389: gap of 100 bp
13390
15335: contig of 1946 bp in length
15336
15435: gap of 100 bp
15436
16250: contig of 815 bp in length
16251
16350: gap of 100 bp
16351
1669: contig of 1319 bp in length
1670
17769: gap of 100 bp
17770
19907: contig of 2138 bp in length
19908
20007: gap of 100 bp
21470: contig of 1463 bp in length
21471
21571: gap of 100 bp
21572
23086: contig of 1516 bp in length
23087
23186: gap of 100 bp
24723: contig of 1537 bp in length
24724
24823: gap of 100 bp
27446: contig of 2923 bp in length
27447
27846: gap of 100 bp
27847
29873: contig of 2027 bp in length
29874
29973: gap of 100 bp
29974
33212: contig of 3239 bp in length
33213
33312: gap of 100 bp
33313
61733: contig of 28420 bp in length
61733
61833: gap of 100 bp
61833
63792: contig of 1960 bp in length
63792
63793: gap of 100 bp
63793
63892: gap of 100 bp
63893
67231: contig of 3339 bp in length
67232
67232: gap of 100 bp
67232
70922: contig of 3591 bp in length
70922
71023: gap of 100 bp
71023
76930: contig of 5908 bp in length
76930
76931: gap of 100 bp
76931
82623: contig of 553 bp in length
82624
82723: gap of 100 bp
82724
87875: contig of 5152 bp in length
87876
87975: gap of 100 bp
87976
93774: contig of 5799 bp in length
93775
93874: gap of 100 bp
93875
93875: contig of 10331 bp in length
93875
104206: gap of 100 bp
104206
104306: contig of 14799 bp in length
104306
119104: gap of 100 bp
119104
119204: gap of 100 bp
119205
132299: contig of 13095 bp in length
132300
132300: gap of 100 bp
132300
148970: contig of 16571 bp in length
148970
149070: gap of 100 bp
149070
149071: gap of 100 bp
149071
155714: contig of 16644 bp in length
155714
155814: gap of 100 bp
155815
16368: contig of 17355 bp in length
16368
183470: gap of 100 bp
183470
210291: contig of 26822 bp in length
210292
210391: gap of 100 bp
210392
212848: contig of 2457 bp in length.
Location/Qualifiers

```
source
1. .212848
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_lib="RP23-300K5"
/clone_1ib="RP23-300K5"
misc_feature
1. .641
/note="assembly_fragment"
742. .1427
/note="assembly_fragment"
misc_feature
1528. .3170
/note="assembly_fragment"
misc_feature
3271. .4396
/note="assembly_fragment"
misc_feature
4497. .5909
/note="assembly_fragment"
misc_feature
6010. .7583
/note="assembly_fragment"
misc_feature
7684. .8975
/note="assembly_fragment"
misc_feature
9076. .10257
/note="assembly_fragment"
misc_feature
10358. .11608
/note="assembly_fragment"
misc_feature
11709. .13289
/note="assembly_fragment"
misc_feature
13390. .15335
/note="assembly_fragment"
misc_feature
15436. .16250
/note="assembly_fragment"
misc_feature
16351. .17669
/note="assembly_fragment"
```

```
Query Match 28.2%; Score 360; DB 2; Length 212848;
Best Local Similarity 91.7%; Pred. No. 2,1e-64;
Matches 425; Conservative 0; Mismatches 21; Indels 18; Gaps 4;

QY 336 AGCCAGCCCGAGCCGATCGGATCGTCTCAACCCCGGCGCAAGCCGCGCGCGGAG 395
DB 93975 AGCCAGCCCGAGCCGATCGGATCGTCTCAACCCCGGCGCAAGCCGCGCGGAG 93994
QY 396 ATCCCGCGATCTGTGAGACCGTACCCCGGCTCTGCTGCTGAGCTTCTGTGCTCTC 455
DB 93935 ATCCCGCGATCTGTGAGACCGTACCCCGGCTCTGCTGCTGAGCTTCTGTGCTCTC 93990
QY 456 CTCTCACTGAGAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 515
DB 93991 CTCTCACTGAGAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94049
QY 516 ATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 575
DB 94050 ATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94109
QY 576 GCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 635
DB 94110 GCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94169
QY 636 ACCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 695
DB 94170 ACCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94229
QY 696 ATGTCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 755
DB 94230 GTGTCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94276
QY 756 CTCCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 803
DB 94277 CTCCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94324

RESULT 11
AY191416 1275 bp mRNA linear PRI 12-MAR-2003
LOCUS Homo sapiens zygote arrest 1 (ZAR1) mRNA, complete cds.
DEFINITION
```

```
ACCESSION AY191416
VERSION AY191416.1 GI:27808691
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1275)
AUTHORS Wu,X., Viveiros,M.M., Eppig,J.J., Bai,Y., Fitzpatrick,S.L. and
Matzuk,M.M.
TITLE Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for
the oocyte-to-embryo transition
Natl. Genet. 33 (2), 187-191 (2003)
MEDLINE 22447938
PUBMED 12539046
REFERENCE 2 (bases 1 to 1275)
AUTHORS Wu,X., Bai,Y. and Matzuk,M.M.
TITLE Direct Submission
Submitted (04-DEC-2002) Pathology, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
FEATURES
source
1. .1275
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4p12"
1. .1275
/gene="ZAR1"
1. .1275
/gene="ZAR1"
/codon_start=1
/product="zygote arrest 1"
/protein_id="AA024707.1"
/db_xref="GI:27808692"
/translation="MAALDEVLVDGVFPACPPCSRYRYPATKGAAGSGMQRC
RGICPASPAGASISFPFCGRILTAAYFDSYOSERLMAIAVYGPGLPARAG
SCDVAVVSPRIDAIVOCISLGRRTIORAPADPSAPGAPAGSGSFSQOBSRL
EGSPVONGARPRPRFRTVAVYSPILARLTALRPPGAPAGSGSFSQOBSRPA
RLQSPFEGEYWTIKAFRRPOSDDDEHQAIVRASWEDPADGPLPREABEGAAPRS
ALSPFGPSPAGVARDGDREAVAGSGSPSPEDGKRLFPUECKRYGHCD
CNIRMSAVYACVQNTKVFQFCRTQCKSYNRYVEDITTCQSKQTRCSCHVKLKH
VDPARPHRDLQCRCKKRLSDSTFSFKYII"
```

```
ORIGIN
Query Match 27.6%; Score 352; DB 9; Length 1275;
Best Local Similarity 64.0%; Pred. No. 1e-62;
Matches 708; Conservative 0; Mismatches 280; Indels 118; Gaps 7;

QY 126 CGCGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 185
DB 171 CGCGGCTCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 230
QY 186 CCACCAAGCGGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 229
DB 231 CTACCAAGCGGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 290
QY 230 --CCGGTCTGTGAGAGCGGCTGACGCTGCGGTGAGAGTGAACCGCGCGAGAGCTTC 287
DB 291 CGCGCGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 350
QY 288 GGTGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 347
DB 351 GGTGCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 410
QY 348 CGCGCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 363
DB 411 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
QY 364 -----TGTCAACCCCGTGGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 413
DB 471 TCGAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 530
```


* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1
642 1427: gap of 100 bp
742 1427: contig of 686 bp in length
1428 1527: gap of 100 bp
1528 3170: contig of 1643 bp in length
3171 3270: gap of 100 bp
3270 4396: contig of 1126 bp in length
4397 4496: gap of 100 bp
4497 5909: contig of 1413 bp in length
5910 6009: gap of 100 bp
6010 7583: contig of 1574 bp in length
7584 7684: gap of 100 bp
7684 8975: contig of 1292 bp in length
8976 9075: gap of 100 bp
9076 10257: contig of 1182 bp in length
10258 10358: gap of 100 bp
10358 11608: contig of 1251 bp in length
11609 11708: gap of 100 bp
11708 13289: contig of 1581 bp in length
13290 13389: gap of 100 bp
13389 15335: contig of 1946 bp in length
15336 15435: gap of 100 bp
15436 16250: contig of 815 bp in length
16251 17659: contig of 1319 bp in length
17660 17769: gap of 100 bp
17770 19907: contig of 2138 bp in length
19908 20007: gap of 100 bp
20008 21470: contig of 1463 bp in length
21471 21570: gap of 100 bp
21571 23086: contig of 1516 bp in length
23087 23187: gap of 100 bp
23187 24724: contig of 1537 bp in length
24725 24823: gap of 100 bp
24824 27746: contig of 2923 bp in length
27747 27846: gap of 100 bp
27847 29873: contig of 2027 bp in length
29874 29973: gap of 100 bp
29974 33212: contig of 3239 bp in length
33213 33312: gap of 100 bp
33313 61733: contig of 28420 bp in length
61734 61833: gap of 100 bp
61834 63792: contig of 1960 bp in length
63793 63892: gap of 100 bp
63893 67231: contig of 3339 bp in length
67232 67331: gap of 100 bp
67332 70923: contig of 3591 bp in length
70924 71023: gap of 100 bp
71024 76930: contig of 5908 bp in length
76931 77030: gap of 100 bp
77031 82623: contig of 5593 bp in length
82624 82723: gap of 100 bp
82724 87875: contig of 5152 bp in length
87876 87975: gap of 100 bp
87976 93774: contig of 5799 bp in length
93775 93874: gap of 100 bp
93875 104205: contig of 10331 bp in length
104206 104305: gap of 100 bp
104306 119104: contig of 14799 bp in length
119105 119204: gap of 100 bp
119205 132299: contig of 13095 bp in length
132300 132399: gap of 100 bp
132400 148970: contig of 16571 bp in length
148971 149070: gap of 100 bp
149071 155714: contig of 16644 bp in length
155715 165815: gap of 100 bp
165816 183369: contig of 17555 bp in length

* 183370 183469: gap of 100 bp
* 183470 210291: contig of 26822 bp in length
* 210292 210391: gap of 100 bp
* 210392 212848: contig of 2457 bp in length.
Location/Qualifiers

FEATURES

source

1..212848

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone_lib="RPCT-23 Female Mouse BAC"

1..641

misc_feature

1..641

/note="assembly_fragment"

742..1427

/note="assembly_fragment"

1528..3170

/note="assembly_fragment"

misc_feature

3271..4396

/note="assembly_fragment"

4497..5909

/note="assembly_fragment"

6010..7583

/note="assembly_fragment"

misc_feature

7684..8975

/note="assembly_fragment"

9076..10257

/note="assembly_fragment"

10358..11608

/note="assembly_fragment"

misc_feature

11709..13289

/note="assembly_fragment"

13390..15335

/note="assembly_fragment"

15436..16250

/note="assembly_fragment"

misc_feature

16351..17659

/note="assembly_fragment"

17660..19907

/note="assembly_fragment"

19908..20007

/note="assembly_fragment"

misc_feature

20008..21470

/note="assembly_fragment"

21471..23086

/note="assembly_fragment"

23087..23187

/note="assembly_fragment"

misc_feature

23188..24724

/note="assembly_fragment"

24725..24823

/note="assembly_fragment"

24824..27746

/note="assembly_fragment"

misc_feature

27747..27846

/note="assembly_fragment"

27847..29873

/note="assembly_fragment"

29874..29973

/note="assembly_fragment"

misc_feature

29974..33212

/note="assembly_fragment"

33213..61733

/note="assembly_fragment"

61734..61833

/note="assembly_fragment"

misc_feature

61834..63792

/note="assembly_fragment"

63793..67231

/note="assembly_fragment"

67232..67331

/note="assembly_fragment"

misc_feature

67332..70923

/note="assembly_fragment"

70924..71023

/note="assembly_fragment"

71024..76930

/note="assembly_fragment"

misc_feature

76931..77030

/note="assembly_fragment"

77031..82623

/note="assembly_fragment"

82624..82723

/note="assembly_fragment"

misc_feature

82724..87875

/note="assembly_fragment"

87876..87975

/note="assembly_fragment"

87976..93774

/note="assembly_fragment"

misc_feature

93775..104205

/note="assembly_fragment"

104206..104305

/note="assembly_fragment"

104306..119104

/note="assembly_fragment"

misc_feature

119105..132299

/note="assembly_fragment"

132300..148970

/note="assembly_fragment"

148971..149070

/note="assembly_fragment"

misc_feature

149071..155714

/note="assembly_fragment"

155715..165815

/note="assembly_fragment"

165816..183369

/note="assembly_fragment"

misc_feature

183370..183469

/note="assembly_fragment"

183470..210291

/note="assembly_fragment"

210292..210391

/note="assembly_fragment"

misc_feature

210392..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"

misc_feature

212849..212848

/note="assembly_fragment"

212849..212848

/note="assembly_fragment"</


```

RESULT 13
LOCUS AY283176 1052 bp mRNA linear VRT 21-AUG-2003
DEFINITION Xenopus laevis zygote arrest 1 (Zarl) mRNA, complete cds.
ACCESSION AY283176
VERSION AY283176.1 GI:30908934
KEYWORDS
SOURCE
ORGANISM Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodidae; Xenopus.
REFERENCE
AUTHORS Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.
TITLE Zygote arrest 1 (zar1) is an evolutionarily conserved gene
expressed in vertebrate ovaries
JOURNAL Biol. Reprod. 69 (3), 861-867 (2003)
MEDLINE 22811438
PUBMED 12773403
REFERENCE
AUTHORS Wu, X. and Matzuk, M.M.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
FEATURES
source
1..1052
location/Qualifiers
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
1..1052
/gene="Zarl"
24..911
/note="maternal factor"
/codon_start=1
/product="zygote arrest 1"
/protein_id="AAP37038.1"
/db_xref="GI:30908935"
/translation="MYPAVNPYSYRYLNPENKMSMRKNKYLAISYDGTGYCDNYORA
OLKALISQVNPNTLPRLCRANTRDVGVNPDADVOCISLGPRTLLRRPALRPP
PEGSPASPTKTVRFPPTTAVSVAAGRIAPRODCVNLSEKGAIVRESGRGE
GKGDGEIKQMKMDKDEBEAPADOTRPFQLEQRKGYTHKCNITMESAYWCV
QENIKYFQFCRQCSINPVRVEDIMCQSKOTRCACPVKLRHVDPRPHRDLGC
RCGRKRLSCDSTSFXYI"
ORIGIN
Query Match 16.1%; Score 205.8; DB 5; Length 1052;
Best Local Similarity 80.1%; Pred. No. 2.4e-32;
Matches 254; Conservative 0; Mismatches 62; Indels 1; Gaps 1;
QY 796 TTCAGTCTTGAAGCAGAAAGTACGGCTACTATCACTGCAAGAGTCAAAATCCGGTGG 855
DB 594 TTCAGTCTTGAAGCAGAAAGTACGGCTACTATCACTGCAAGAGTCAAAATCCGGTGG 653
QY 856 GAGAGCCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 915
DB 654 GAGAGCCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712
QY 916 CCGAGTGTGAGAAATCTCAACACCTTACAGAGTGAAGACATCACTGTCAAGTTG 975
DB 713 CAGGACATGTCAGAAATCTATATATCCCTACCGTGGAGAACATCATGTGTCAAGCTG 772
QY 976 TAAAGAACTAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035
DB 773 CAGGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832
QY 1036 TCGGCAAGACTTGTGTGGAGATGCAAGCAACGCTGTCTCTGGAGACGACTTTCAG 1095
DB 833 CCGGCAAGACTTGTGTGGAGATGCAAGCAACGCTGTCTCTGGAGACGACTTTCAG 892
QY 1096 CTTCAATATCATCTTT 1112

```

```

DB 893 CTTCAAGTATATCATCTT 909
RESULT 14
LOCUS AY283177 963 bp mRNA linear VRT 21-AUG-2003
DEFINITION Takifugu rubripes zygote arrest 1 (Zarl) mRNA, complete cds.
ACCESSION AY283177
VERSION AY283177.1 GI:30908936
KEYWORDS
SOURCE
ORGANISM Takifugu rubripes (Pugu rubripes)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthopterygii; Acanthopterygii; Perciformes; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
REFERENCE
AUTHORS Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.
TITLE Zygote arrest 1 (zar1) is an evolutionarily conserved gene
expressed in vertebrate ovaries
JOURNAL Biol. Reprod. 69 (3), 861-867 (2003)
MEDLINE 22811438
PUBMED 12773403
REFERENCE
AUTHORS Wu, X. and Matzuk, M.M.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
FEATURES
source
1..963
location/Qualifiers
/organism="Takifugu rubripes"
/mol_type="mRNA"
/db_xref="taxon:31033"
1..963
/gene="Zarl"
1..963
/note="maternal factor"
/codon_start=1
/product="zygote arrest 1"
/protein_id="AAP37039.1"
/db_xref="GI:30908937"
/translation="MATYCDPEVDVSYYSSINPMPGRYPHRRDAGMKYKYSLYSGDT
SEAFSNOQRAQLKSLISQINPKLPRLRKATKQVAVQVQKRVASVQCS:GPRTLV
VKRLRRRRKLNPGPGTPOKTEBGEVYPRLAYSSIAFESVTSFVETSKDPRALF
AAAEELPQEPQEPQCGENQAGEYETNANLPEGRKQSDQADQADAEBSKGAARVFOR
LEOKYGVHCRKCNLRMSAYWCVQGNKYVFPQFCRCKCKDNPVRVEDITGVHVN
KARCACAETQHVDPKRPHRDLGCRKGRKRLSCDSTSFXYIY"
ORIGIN
Query Match 15.3%; Score 194.8; DB 5; Length 963;
Best Local Similarity 75.4%; Pred. No. 4.7e-30;
Matches 255; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
QY 776 AGGACAAAGAGCGCTGCTTCCAGTCTTAGAGCAGAGTACGGCTACTATCACTGCA 835
DB 626 AGGACAAAGCGCTGCTTCCAGTCTTAGAGCAGAGTACGGCTACTATCACTGCA 685
QY 626 AGGACAAAGCGCTGCTTCCAGTCTTAGAGCAGAGTACGGCTACTATCACTGCA 685
DB 626 AGGACAAAGCGCTGCTTCCAGTCTTAGAGCAGAGTACGGCTACTATCACTGCA 685
QY 836 AGGACTGCAAAATCCGGTGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
DB 686 GAGAAATGCAACCTGAGATGGAGAGCGCGTACGTTTGGTGGTTCAAGGCACTAAAGG 745
QY 896 TGTACTTCAAAAGCTTCTGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 955
DB 746 T-TTACTTCAAGCAGTTCTGTGAGAAATGCCAAAAGACTTTAAACCGTACCGGTAGAG 804
QY 956 GACATCACTGTCAAAGTTGTAAAAGAACTAGATGTGCTGCCAGTCAATTTGGCCAC 1015
DB 805 GACATCACTGTCAAAGTTGTAAAAGAACTAGATGTGCTGCCAGTCAATTTGGCCAC 864
QY 1016 GTGACCTTAAAGCCCCCATCGGCAAGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1075
DB 865 GTTACCCAAAGAGCCGCCACAGGACAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 924

```

QY 1076 TCCTGCGACGACCTTCACTTCAATATCATTTA 1113
 Db 925 TCCTGCGACGACCTTCACTTCAATATCATTTA 962

RESULT 15
 LOCUS

AY283178 1084 bp mRNA linear VRT 21-AUG-2003
 DEFINITION Danio rerio zygote arrest 1 (zar1) mRNA, complete cds.
 ACCESSION AY283178
 VERSION AY283178.1 GI:30908938

KEYWORDS

Danio rerio (zebrafish)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.

REFERENCE

Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.
 Zygote arrest 1 (zar1) is an evolutionarily conserved gene
 expressed in vertebrate ovaries
 Biol. Reprod. 69 (3), 861-867 (2003)

JOURNAL

22811438

MEDLINE

12773403

REFERENCE

2 (bases 1 to 1084)

AUTHORS

Wu, X. and Matzuk, M.M.

TITLE

Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

JOURNAL

Location/Qualifiers

1..1084
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 1..1084
 /gene="zar1"
 42..1031
 /gene="zar1"
 /note="maternal factor"
 /codon_start=1
 /product="zygote arrest 1"
 /protein_id="AAP37040.1"
 /db_xref="GI:30908939"

ORIGIN

Query Match 15.2%; Score 194.6; DB 5; Length 1084;
 Best Local Similarity 75.1%; Pred. No. 5.2e-30;
 Matches 256; Conservative 0; Mismatches 84; Indels 1; Gaps 1;

774 GCAGAGCAAGAGCGCGCTGCTTCCAGTCTTAGAGCAGAAAGTACGCTACTATCATCTG 833
 692 GAAGTCCAGAGCTCGTGTGATTTCAAGCTTTGAGCAGAAAGTATGATTTATCATCTG 751
 834 CAAGAGCTGCAAAATCCGCTGAGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGT 893
 752 CAAGAGCTGCAAAATCCGCTGAGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGT 811
 894 GGTGTACTTCAACAAGTCTGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 953
 812 GGT-TTATTTCAGAGCTGT 870
 954 AGGACATCACTGTCTCAAGTTGTAAGAACTAGATGTGCTGCCAGTCAATTTGCGC 1013
 871 AGGACATGAGCATGTGAGCTTGCAGAAAGCTGTCGACATGTTCTGTCAAGTCGCGTC 930
 1014 ACGTGACCTTAAGCCCCCATCGGCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1073

Db 931 ACCTGAGACCCCAAGAACCCCATCGGACGATCTGTGCGGCCCGCTGTAAGGCAAGCGTC 990
 QY 1074 TGCTGTGACGACGACCTTCACTTCAATATCATTTAG 1114
 Db 991 TGCTGTGACGACGACCTTCACTTCAATATCATTTAG 1031

Search completed: April 6, 2004, 19:34:21
 Job time: 5131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 17:58:29 ; Search time 528 Seconds

9064.163 Million cell updates/sec

Title: US-09-830-810A-1
 Page: 1377

Sequence: 1 aaagcgggcgaaggcgcgga.....acaaadadaadadaaaaaa 1277

Gapop 1.0.0 , Gapext 1.0.0

Searched: 2470632 seqs, 1873875610 residues

Total number of hits satisfying chosen parameters: 4941264

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database

```

1: /cgn2_6/prodataa2/pubnpna/us97_PUBCOMB.seq:*
2: /cgn2_6/prodataa2/pubnpna/PCPT_NEW_PUB.seq:*
3: /cgn2_6/prodataa2/pubnpna/us06_NEW_PUB.seq:*
4: /cgn2_6/prodataa2/pubnpna/us06_PUBCOMB.seq:*
5: /cgn2_6/prodataa2/pubnpna/us07_NEW_PUB.seq:*
6: /cgn2_6/prodataa2/pubnpna/PCPTs_PUBCOMB.seq:*
7: /cgn2_6/prodataa2/pubnpna/us08_NEW_PUB.seq:*
8: /cgn2_6/prodataa2/pubnpna/us08_PUBCOMB.seq:*
9: /cgn2_6/prodataa2/pubnpna/us09a_PUBCOMB.seq:*
10: /cgn2_6/prodataa2/pubnpna/us09b_PUBCOMB.seq:*
11: /cgn2_6/prodataa2/pubnpna/us09c_PUBCOMB.seq:*
12: /cgn2_6/prodataa2/pubnpna/us09_NEW_PUB.seq:*
13: /cgn2_6/prodataa2/pubnpna/us10a_PUBCOMB.seq:*
14: /cgn2_6/prodataa2/pubnpna/us10c_PUBCOMB.seq:*
15: /cgn2_6/prodataa2/pubnpna/us10_NEW_PUB.seq:*
16: /cgn2_6/prodataa2/pubnpna/us60_NEW_PUB.seq:*
17: /cgn2_6/prodataa2/pubnpna/us60_PUBCOMB.seq:*
18: /cgn2_6/prodataa2/pubnpna/us60_PUBCOMB.seq:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1277	100.0	1277	9	US-09-844-864-1	Sequence 1, Appl
2	799.4	52.6	814	9	US-09-844-864-18	Sequence 18, Appl
3	752.2	58.9	809	9	US-09-844-864-22	Sequence 22, Appl
4	273.6	21.4	335	9	US-09-844-864-21	Sequence 21, Appl
5	260.6	20.4	375	9	US-09-844-864-25	Sequence 25, Appl
6	98.8	7.7	123	9	US-09-844-864-19	Sequence 19, Appl
7	98.8	7.7	123	9	US-09-844-864-23	Sequence 23, Appl
8	72	5.6	105	9	US-09-844-864-20	Sequence 20, Appl
9	70.4	5.5	105	9	US-09-844-864-24	Sequence 24, Appl
10	54.2	4.2	1500	12	US-10-413-6598-1133	Sequence 1133, Appl
11	54.2	4.2	1500	15	US-10-377-7804-680	Sequence 680, Appl
12	51.6	4.0	987	15	US-10-084-8464-99	Sequence 99, Appl
13	51.6	4.0	59816	15	US-10-084-8464-1	Sequence 1, Appl
14	51.6	4.0	59816	15	US-10-084-8464-2	Sequence 2, Appl
15	51	4.0	1185	14	US-10-156-761-6879	Sequence 6879, Appl

16	51	4.0	9025568	14	US-10-156-761-1	Sequence 1, App1
C 17	50.2	3.9	1673	15	US-10-112-6998-1140	Sequence 1140, Ap
C 18	50.2	3.9	1673	15	US-10-193-7800-687	Sequence 687, App
C 19	49.6	3.9	400	14	US-10-1374-002-174	Sequence 174, App
C 20	49.6	3.9	400	14	US-10-084-843-179	Sequence 179, App
C 21	49.6	3.9	14800	9	US-09-954-456-1601	Sequence 1601, Ap
C 22	49.6	3.9	14800	14	US-10-269-909-61	Sequence 61, App1
C 23	49.6	3.9	14835	14	US-10-240-965-113	Sequence 113, App
C 24	49	3.8	7175	15	US-10-375-253-13	Sequence 13, App1
C 25	49	3.8	7177	13	US-10-033-026-7	Sequence 7, App1
C 26	49	3.8	7362	15	US-10-975-253-11	Sequence 11, App1
C 27	49	3.8	7364	9	US-09-954-456-11179	Sequence 1179, Ap
C 28	49	3.8	7364	13	US-10-033-026-5	Sequence 5, App1
C 29	49	3.8	7376	13	US-10-033-026-3	Sequence 3, App1
C 30	48.6	3.8	633	15	US-10-027-632-154723	Sequence 154723, App
C 31	48.6	3.8	633	15	US-10-027-632-154724	Sequence 154724, App
C 32	48.6	3.8	633	15	US-10-027-632-154725	Sequence 154725, App
C 33	47.8	3.7	809	15	US-10-027-632-165400	Sequence 165400, App
C 34	47.8	3.7	809	15	US-10-027-632-165401	Sequence 165401, App
C 35	47.6	3.7	5944	14	US-10-311-455-228	Sequence 226, App
C 36	47	3.7	1026	12	US-10-085-782A-53529	Sequence 53529, A
C 37	47	3.7	1026	15	US-10-242-535A-53529	Sequence 53529, A
C 38	47	3.7	4266	14	US-10-156-761-5083	Sequence 5083, Ap
C 39	46.8	3.7	536	16	US-10-138-110-119	Sequence 119, App
C 40	46.6	3.6	745	8	US-08-736-019-163	Sequence 163, App
C 41	46.6	3.6	745	10	US-09-366-886-51	Sequence 51, App1
C 42	46.6	3.6	745	10	US-09-373-658-70	Sequence 70, App1
C 43	46.6	3.6	745	11	US-09-989-668-104	Sequence 104, App1
C 44	46.6	3.6	1006	9	US-09-795-668-104	Sequence 104, App
C 45	46.6	3.6	1006	9	US-09-795-668-104	Sequence 104, App

ALIGNMENTS

```

; RESULT 1
; US-09-844-864-1
; Sequence 1, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: Ovary Specific Genes and Proteins
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-844-864-1

```

Query Match	100.0%;	Score 1277;	DB 9;	Length 1277;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1277; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY 1 AAGGCGGGGAGAGCGGGGAGCGACCCCATGTTCCCGGCGAGCAGTCGTCACCCCTTGCCCG 60
Db 1 AAGCGCGGAGAGCGGGGAGCGACCCCATGTTCCCGCGAGCAGTCGTCACCCCTTGCCCG 60
QY 61 CATCTTATCCGACGAGCCAAAGCGGGGATGCGTGAAGTTCGGAACCAAGGGCTGC 12
Db 61 CATCTTATCCGACGAGCCAAAGCGGGGATGCGTGAAGTTCGGAACCAAGGGCTGC 12
QY 121 CGACCGGCGCCCGCTCTCTCTCCCGGCTACAGCAGCTATGGCGCGGAGTACGTC 18

Db 121 GAGCCGCGCCCTCTCTCTCCCGGCTACAGACGCTCATGGCCGGAGTACGTC 180
QY 181 GACAGCCACAGCGGGGCAAGCTCATGCGCTCTGCGGATGGGTCGCCGTCGCTC 240
Db 181 GACAGCCACAGCGGGGCAAGCTCATGCGCTCTGCGGATGGGTCGCCGTCGCTC 240
QY 241 AGCAGCCGTCGCGTGGAGTGAACCGCGCGCGAGAGCGCTGGTGGAGTGTCA 300
Db 241 AGCAGCCGTCGCGTGGAGTGAACCGCGCGCGAGAGCGCTGGTGGAGTGTCA 300
QY 301 CTGGGCGCGGCAAGCTGACGCTGCGAGGAGTCCGAGCCAGCGCGCGATCGCT 360
Db 301 CTGGGCGCGGCAAGCTGACGCTGCGAGGAGTCCGAGCCAGCGCGCGATCGCT 360
QY 361 TCCGTCAACCCCGTGGGCGACGCGCGCGCGAGATCCCGGATCTGGAGACCGTA 420
Db 361 TCCGTCAACCCCGTGGGCGACGCGCGCGCGAGATCCCGGATCTGGAGACCGTA 420
QY 421 GCGCCGTTCTGCTCGGTGACCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 421 GCGCCGTTCTGCTCGGTGACCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 481 AGCAGACACCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 AGCAGACACCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 CCGAGAGAGGTGGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 CCGAGAGAGGTGGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 CAGGCTGACAGGAG 660
Db 601 CAGGCTGACAGGAG 660
QY 661 GCGGCGATGAGCTGAGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 GCGGCGATGAGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GACCCGCGGTGATTCGAGATGCGCTCGAGACAGAGCTCCCGGAGAGAGAGAG 780
Db 721 GACCCGCGGTGATTCGAGATGCGCTCGAGACAGAGCTCCCGGAGAGAGAGAG 780
QY 781 AAGGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 AAGGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 TGCAGAAATCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 TGCAGAAATCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 CTTCAGAAAGTTCGCGAGTGTGAGAAATCTTAACAACCTTCAACAGTGTGAGAG 960
Db 901 CTTCAGAAAGTTCGCGAGTGTGAGAAATCTTAACAACCTTCAACAGTGTGAGAG 960
QY 961 CACCTGTCAAAGTGTGAGAAAGTGTGAGTGTGCTGCGAGTCAAGTTCGCAAGTGA 1020
Db 961 CACCTGTCAAAGTGTGAGAAAGTGTGAGTGTGCTGCGAGTCAAGTTCGCAAGTGA 1020
QY 1021 CCTTAACGCGCCCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 CCTTAACGCGCCCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 CGACAGACCTTCAGCTTCAAAATACATCATTTAGAGAGTGAAGAAAGTTCGCTAGA 1140
Db 1081 CGACAGACCTTCAGCTTCAAAATACATCATTTAGAGAGTGAAGAAAGTTCGCTAGA 1140
QY 1141 TGGGCGTAAATGAGATGAGACAGTGTTCCTCTCTCTCTCTCTCTCTCTCTCT 1200
Db 1141 TGGGCGTAAATGAGATGAGACAGTGTTCCTCTCTCTCTCTCTCTCTCTCTCT 1200
QY 1201 TCTTCATGACAGACAGTGTCTTACTGAGATTAAGAGCTGTGATTAAGGATTCGAGACA 1260
Db 1201 TCTTCATGACAGACAGTGTCTTACTGAGATTAAGAGCTGTGATTAAGGATTCGAGACA 1260

QY 1261 AAAAAAAAAAAAAAAAAA 1277
Db 1261 AAAAAAAAAAAAAAAAAA 1277

RESULT 2
US-09-844-864-18
; Sequence 18, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheung
; APPLICANT: Xu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 814
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-18

Query Match 62.6%; Score 799.4; DB 9; Length 814;
Best Local Similarity 99.9%; Pred. No. 6,2e-225;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAGCGGAG 62
Db 1 GCGCGGAG 60
QY 63 TCCATTATCCGAGGACCAACCAAGCCGAGAGTGGCTGAGAGTCCGAGAGAGAGAG 122
Db 61 TCCATTATCCGAGGACCAACCAAGCCGAGAGTGGCTGAGAGTCCGAGAGAGAGAG 120
QY 123 ACCCGCGCCCT 182
Db 121 ACCCGCGCCCT 180
QY 183 CAGGACACAGCGGAGCAGCTCATGAGCCTGCTGTGCGAGATGGTCCCGGTCAG 242
Db 181 CAGGACACAGCGGAGCAGCTCATGAGCCTGCTGTGCGAGATGGTCCCGGTCAG 240
QY 243 CAGCGGAG 302
Db 241 CAGCGGAG 300
QY 303 CCGGCGCGGACGCTGAGCTGAGCTGAGAGTGGCCAGAGCAGCCGAGTGGGTTTC 362
Db 301 CCGGCGCGGACGCTGAGCTGAGCTGAGAGTGGCCAGAGCAGCCGAGTGGGTTTC 360
QY 363 CTGTCAACCCCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 422
Db 361 CTGTCAACCCCGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 423 CCGGTTCTGCTCGGTGACCTTCTGTGAGCTCTCTCTCTCTCTCTCTCTCTCTCT 482
Db 421 CCGGTTCTGCTCGGTGACCTTCTGTGAGCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 483 GCGAGACACCCACGAG 542
Db 481 GCGAGACACCCACGAG 540
QY 543 GAGAGAGTGGCGGAG 602
Db 541 GAGAGAGTGGCGGAG 600

QY 603 GGCTGACAGGCGCGGCTGGAGAGACAGCCACACCGAGAGACCGGAACAGTGTGC 662
 Db 601 GGCTGACAGGCGCGGCTGGAGAGACAGCCACACCGAGAGACCGGAACAGTGTGC 660
 QY 663 GAGCATGAGTCTGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
 Db 661 GAGCATGAGTCTGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 QY 723 CCCCAGTATTGAGATGCCCCCTCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
 Db 721 CCCCAGTATTGAGATGCCCCCTCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 783 GAGAGCGCTGCGTTTCCAGTT 803
 Db 781 GAGAGCGCTGCGTTTCCAGTT 801

RESULT 3
 US-09-844-864-22
 ; Sequence 22, Application US/09844864
 ; Patent No. US20020042926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Matzlik, Martin
 ; APPLICANT: Ren, Yongsheung
 ; APPLICANT: Wu, Xuemei
 ; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 ; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
 ; CURRENT APPLICATION NUMBER: US/09/844,864
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/106,020
 ; PRIOR FILING DATE: 1998-10-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/25209
 ; PRIOR FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 809
 ; TYPE: DNA
 ; ORGANISM: mus musculus
 US-09-844-864-22

Query Match 58.9%; Score 752.2; DB 9; Length 809;
 Best Local Similarity 97.1%; Pred. No. 5e-211;
 Matches 778; Conservative 0; Mismatches 18; Indels 5; Gaps 1;
 QY 3 GGCAGGCGAGGCGCGGAG 62
 Db 1 GGCAGGCGAGGCGCGGAG 60
 QY 63 TCCTTATTCGAGAGCGACCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
 Db 61 TCCTTATTCGAGAGCGACCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 123 ACCGCGCGCGCGCT 182
 Db 121 ACCGCGCGCGCGCT 180
 QY 183 CAGCAGCAGGCGGAG 242
 Db 181 CAGCAGCAGGCGGAG 240
 QY 243 CAGCAGCAGGCGGAG 302
 Db 241 CAGCAGCAGGCGGAG 300
 QY 303 CCGAGCGCGAG 362
 Db 301 CCGAGCGCGAG 360
 QY 363 CTGTCAGACCCGCGAG 422
 Db 361 CTGTCAGACCCGCGAG 420

QY 423 CCGGTTCTGTCGCTGAGACCTTCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
 Db 421 CCGGTTCTGTCGCTGAGACCTTCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 QY 483 GAGAGACACCCAG 542
 Db 481 GAGAGACACCCAG 540
 QY 543 GAGAGAGTGGCGGAG 602
 Db 541 GAGAGAGTGGCGGAG 600
 QY 603 GAGTGAAGGAG 662
 Db 601 GAGTGAAGGAG 660
 QY 663 GAGCATGAGTCTGAG 722
 Db 661 GAGCATGAGTCTGAG 720
 QY 723 CCCCAGTATTGAGATGCCCCCTCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
 Db 721 CCCCAGTATTGAGATGCCCCCTCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
 QY 783 GAGAGCGCTGCGTTTCCAGTT 803
 Db 776 GAGAGCGCTGCGTTTCCAGTT 796

RESULT 4
 US-09-844-864-21
 ; Sequence 21, Application US/09844864
 ; Patent No. US20020042926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Matzlik, Martin
 ; APPLICANT: Ren, Yongsheung
 ; APPLICANT: Wu, Xuemei
 ; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 ; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
 ; CURRENT APPLICATION NUMBER: US/09/844,864
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/106,020
 ; PRIOR FILING DATE: 1998-10-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/25209
 ; PRIOR FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 21
 ; LENGTH: 305
 ; TYPE: DNA
 ; ORGANISM: mus musculus
 US-09-844-864-21

Query Match 21.4%; Score 273.6; DB 9; Length 305;
 Best Local Similarity 95.3%; Pred. No. 3.6e-70;
 Matches 282; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 965 TGTCAAAGTTGTAAG 1024
 Db 10 TGTCAAAGTTGTAAG 96
 QY 1025 AAAGCGCGCGAG 1084
 Db 70 AAAGCGCGCGAG 129
 QY 1085 AGACACCTTCAAGCTTCAAAATACATCATTTAGTGAAGTGAAGAGAGAGAGAGAGAG 1144
 Db 130 AGACACCTTCAAGCTTCAAAATACATCATTTAGTGAAGTGAAGAGAGAGAGAGAGAG 189
 QY 1145 GGTATGAG 1204
 Db 190 GGTATGAG 249

QY 1205 CATGACAGACGTGTTACTTGGATTATAAGCCCTGTGAATAAAAGTATTGCCAAACA 1266

Db 250 CATGACAGACGTGTTACTTGGATTATAAGCCCTGTGAATAAAAGTATTGCCAAACA 305

```

RESULT 5
US-09-844-864-25
  Sequence 25, Application US/09844864
  Patent NO. US20020042926A1
  GENERAL INFORMATION:
    APPLICANT: Matzink, Martin
    APPLICANT: Ren, Yongsheng
    APPLICANT: Wu, Xuemei
    TITLE OF INVENTION: Ovary specific genes and proteins
    FILE REFERENCE: P01928US2 / 09807797 / OIA 99-48
    CURRENT APPLICATION NUMBER: US/09/844,864
    CURRENT FILING DATE: 2001-04-27
    PRIOR APPLICATION NUMBER: 60/106,020
    PRIOR FILING DATE: 1998-10-28
    PRIOR APPLICATION NUMBER: PCT/US99/25209
    PRIOR FILING DATE: 1999-10-28
    NUMBER OF SEQ ID NOS: 25
    SOFTWARE: PatentIn version 3.0
    SEQ ID NO 25
      LENGTH: 375
      TYPE: DNA
    ORGANISM: mus musculus
  US-09-844-864-25

```

Query Match	20.4%;	Score 260.6;	DB 9;	Length 375;
Best Local Similarity	96.7%;	Pred. No. 2.7e-66;		
Matches 266;	Conservative	0;	Mismatches 9;	Indels 0

Dy 986 AGATGTGCTGCCAGTCAGATTTCGCACGTGGACCCCAAGCGCCCCCATCGGCAAGAC 1043
Dd 101 AGATGTGCTGCCAGTCAGACCTCGCACGTGAACCTTAGACGGCCCCCATCAGCAAGAC 160

Dy 1046 TTGTGGGAGATGCACGAACAACGCCCTGTCCCTCGCAGACACTTCAAGCTTAATAATC 1105
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 161 TTTGTGAGAGATGCACGAACAACGCCCTGTCCCTCGCAGACACCGTCAGCTTCAATAATC 220

Qy 1106 ATCATTTTGTGAGAGTCCAAAACGTTTCCTAGATGGGGCTAATGGAATGGACAACTGA 1155

Dp 221 ATGATTTTGTGAGAGTCGAAAACGTTTCCTAGATGGGGCTAATGGAATGGACAACTGA 280

Dy 1166 GCTTCTCCCTCTTCACCTCTTCCCCTTCCAAATTCTTCATGACAGACAGTGTTACTTGG 1223

Dd 281 GCCTTCTCCCCCTTCACCTCTTCCCCTTCCAAATTCTTCATGACAGACAGTGTTACTTGG 340

Qy	1226	GATATAAGCCTGTGAATAAAGTATTGCCAACA	1280
Db	341	GATATAAGCCTGTGAATAAAGTATTGCCAACA	375

US-RSULT 6
US-09-844-864-19
Sequence 19, Application US/09844864
Patent No US20070042926A1
GENERAL INFORMATION:
APPLICANT: Matzuk, Martin
APPLICANT: Ren, Yongsheng
APPLICANT: Wu, Xumei
TITLE OF INVENTION: Ovary Specific Genes And Proteins
FILE REFERENCE: P01925052 / US9807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106,020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0

```

; SEQ ID NO 19
; LENGTH: 123
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-19

```

Query Match	7.7%	Score 98.8	DB 9	Length 123
Best Local Similarity	98.0%	Pred. No. 7.6e-19		
Matches 100; Conservative	0	Mismatches 2	Indels 0	Gaps 0

QY 795 TTTCAGTCTTAGAGCAGAGTACGGCTACTACTGTCAAGACTGCAGAAATCCGGTG 85
Db 9 TGTTCAGTCTTAGAGCAGAGTACGGCTACTACTGTCAAGACTGCAGAAATCCGGTG 68

```

QY      855 GGAGAGCGCCTATGTGTGGTGTGTGACGGGACCAAGTAAGT 896
      |||||
Db      69  GGAGAGCGCCTATGTGTGGTGTGTGACGGGACCAAGTAAGT 110

```

```

RESULT 7
US-09-844-864-23
; Sequence 23, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:

```

? APPLICANT: Matzku, Martin
 ? APPLICANT: Ren, Yongsheng
 ? APPLICANT: Wu, Xuemei
 ? TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 ? FILE REFERENCE: P0192US2 / 09607797 / ODA 99-48
 ? CURRENT APPLICATION NUMBER: US/09/844,864
 ? CURRENT PRIORITY NUMBER: 44,250,000

```

1 /
2 / CURRENT FILING DATE: 2001-04-27
3 / PRIOR APPLICATION NUMBER: 60/106,020
4 / PRIOR FILING DATE: 1998-10-28
5 / PRIOR APPLICATION NUMBER: PCT/US99/25209
6 / PRIOR FILING DATE: 1999-10-28
7 / NUMBER OF SEQ ID NOS: 25
8 / SOFTWARE: PatentIn version 3.0

```

```

! SEQ ID NO 23
!
! LENGTH: 123
! TYPE: DNA
! ORGANISM: mus musculus
US-09-844-864-23

```

Query Match	7.7%;	Score 98.8;	DB 9;	Length 123;
Best Local Similarity	98.0%;	Pred. No. 7.8e-19;		
Matches 100;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 795 TTTCAGTCTTAGACAGAATCCGGTAATACTGCAAGACTGCAAAATCCGGTG 854
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 TGTTCAGTTCTTAGACAGAAATCCGGTAATACTGCAAGACTGCAAAATCCGGTG 68

QY 855 GGAGAGCGCCTATGTGTGTGTGTGTCAGGGGCACCAGTAAGT 896
|||
Db 69 GGAGAGCGCCTATGTGTGTGTGTCAGGGGCACCAGTAAGT 110

RESULT 8
US-09-844-864-20
Sequence 20, Application US/09844864
Patent No. US20020042926A1
GENERAL INFORMATION:
APPLICANT: Matczuk, Martin
APPLICANT: Ren, Yongsheng
APPLICANT: Wu, Xuemei
TITLE OF INVENTION: CRYA SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P0192US2 / 09607797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 68/106,020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 105
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-20

Query Match
Best Local Similarity 100.0%; Pred. No. 5.7e-11; Length 105;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 899 TACTTCAACAGTTCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 958
Db 19 TACTTCAACAGTTCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 78

QY 959 ATCAGCTGTCAA 970
Db 79 ATCAGCTGTCAA 90

RESULT 9
US-09-844-864-24
; Sequence 24, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 105
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-24

Query Match
Best Local Similarity 98.6%; Pred. No. 1.7e-10; Length 105;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 899 TACTTCAACAGTTCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 958
Db 19 TACTTCAACAGTTCTGCCGAGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGAC 78

QY 959 ATCAGCTGTCAA 970
Db 79 GTCACCTGTCAA 90

RESULT 10
US-10-412-699B-1133
; Sequence 1133, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Ficomm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Biron, Pierre E.
; APPLICANT: Pineda, Omalta
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
```

```

; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumamoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: M01-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1133
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-412-699B-1133

Query Match
Best Local Similarity 43.4%; Pred. No. 3.7e-05; Length 1500;
Matches 251; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 137 CCTTCTCTCCCGGCTACAGACATGAGTCCCGCGAGTACGTGACAGCCACAGCGGG 196
Db 298 CCTTCTTACACACCCCGACGAGCTCTCGAAGAGGAGTACTACAGACGAGAGCTCCCGG 357

QY 197 CACAGCTATGAGCCCTGTGTGTGGGATGGGTCCCGGTGAGAGCCGTGAGCGTG 256
Db 358 AGAAGAGCGGCGCTTACAGCGCGGAGAGGATGTCTGTGAGAGAGCTTCAAGAGAG 417

QY 257 CGGTGAGGTAAACCCGCGCGGACAGCGCTGCTGTGAGTGTCACTGGGCGCGACGC 316
Db 418 AGAACAAGCTGAGACCGGAGCGGAAGAAGAGCTGGCGCGGAAGCTGAGCGCGC 477

QY 317 TGCAGCTGTGAGGAGGAGCGGAGCGCCCGACAGCGCGGATCGGGTTCGTCAACCCCGTG 376
Db 478 GCGAGGTGCGGTGTGTGTTCAGAAACCGCGCGCGCTGGAACACCAAGAGCTTCAAGC 537

QY 377 GCCAGCGCGGCGCGGAGATCCCGGAGTCTTGGAGACCGTAAACCGGCTTCTGCTCG 436
Db 538 GCGACTTGTGAGCGCTCAAGGCGTGTTCGACGCGCTCCGCGCGACACAGCGCGCTCC 597

QY 437 TGACCTTGTGTGCGCTCTCTCTCTCACTGAGAGTTGGGAGAGGAGGACACACCGCA 496
Db 598 TCAGAGACACACCGGCTTCACTTCAAGTGTGATGTGTGACCGAAGTGTCAAGAGA 657

QY 497 AGGAGAGGAGGAGCGCGCATCTCGGAGAACCCGAGAACCGAGACCGAGAGGTGCGCG 556
Db 658 AGGAGAGGAGGAGGAGGAGCGCGCGCGCGGCGGCTTGAAGTCTCCGAGCTTGGCTCGG 717
```

QY 557 CGAGAAAGCGGTCCCCAGCCGGAAGAGAGCGATGTTCAAGCTGCAGGCGCAG 616
DB 718 CGCGCGAGCTGAAGGTGCGCGTCCCGGACCGCGAGAGACCGCGGCTGAGAGAGCGCGCG 777
QY 617 CCGGTGGAGAC 676
DB 778 CGGCTTCGAGAGAC 837
QY 677 AGCCTGGAGCGAGAC 715
DB 838 GCGCGGAGAGCGCGGTGTGTGACACGAGACCGCAGACTCG 876

RESULT 11
US-10-374-780A-680
; Sequence 680, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Racciliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Biron, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omatra
; APPLICANT: Yu, Guo-Liang
; TITLE OR INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 680
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: *Oryza sativa*
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G409
US-10-374-780A-680

Query Match 4.2%; Score 54.2; DB 15; Length 1500;
Best Local Similarity 43.4%; Pred. No. 3,7e-05;
Matches 251; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 137 CCTCCCTCCCGGCTACAGACAGTCTGCGCGGAGTCTGTCAGACAGCAGCAGCGG 196
DB 298 CTTCTTCCACACCGCCGAGAGCTCTCGAAGAGAGTCTACAGACAGACAGCTCCCG 357

QY 197 CACAGCTCATGAGCCCTGCTGCGGATGGATGCCGTCGCCGTGCTGAGACCGGTACGCTG 256
DB 358 AGAAGAACGGCGGCTTACACGCGCGAGACAGATGATCTGTGAGAGAGAGCTTCAGAGAGG 417
QY 257 CCGTGCAGTGAACCCCGCGCGCGACCGCTCTGTGTGACGTGTTCACTCGGCGCCGACAGC 316
DB 418 AGAACACATCGAGCGCGAGCGAGAGACGAGCTGGCGCGGAACTAGGCGCTGACAGCGC 477
QY 317 TGCAGCTGCAGGGTGCAGACCGACCGCCGACCGCCGATCGGTTCTGTCAACCCCGTG 376
DB 478 GCGAGGTGCGCGTGTGTGTTCAGAACCGCCCGCGCGCTGTGAAACACAGAGCTTCAGC 537
QY 377 GCGACCGCGCGCGCGGAGATCCCGGATCTTGGCAGACCGTACCGCCGTTCTGTCG 436
DB 538 GCGACTTGCAGCGCGCTCAGAGCGCTGTTCGACGCGCTCCGCGCGACACACACCGCTTC 597
QY 437 TGAACCTTGTGCGCTCTCCTCTCACTGAGGTTGGCGGAGAGACAGACACACAGCA 496
DB 598 TCCAGGAAACACACCGCTTCACTCTCACTGATGTGTTGACCGAAGAGCTGCAAGAGA 657
QY 497 AGGAGAGAGGAGCGCGGATCTTGGGAGACCGGAGACCGAGCCGAGAGAGGTGCGG 556
DB 658 AGGAGACGACGACGAGGAGCGCGCGCGCGCTTGAAGTCCCGGCTTGTGCTGCGG 717
QY 557 CGAGAAAGCGGTCCCCAGCGCGGAGAGCGAGGCGGATGTTCAAGCTGCAGGCGCAG 616
DB 718 CGGCGGAGCTAAGGTCCCGTCCCGAGCGCGAGAGACCGCGCTGAGAGAGCGCGCG 777
QY 617 CCGGTGGAGACAGAGCAGACACACACACACACACACACACACACACACACACACAC 676
DB 778 CGGCTTCGAGAGAC 837
QY 677 AGCCTGGAGCGAGAC 715
DB 838 GCGCGGAGAGCGCGGTGTGTGACACGAGACCGCAGACTCG 876

RESULT 12
US-10-084-846A-99
; Sequence 99, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLMEIER, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTOLD, ANDREAS
; TITLE OR INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 99
; LENGTH: 987
; TYPE: DNA
; ORGANISM: *Streptomyces viridochromogenes*
; FEATURE:
; OTHER INFORMATION: aviga data: partial sequence of coding strand 1; nucleotide 1
; OTHER INFORMATION: corresponds to nucleotide 45,341 of coding strand 1.
US-10-084-846A-99

Query Match 4.0%; Score 51.6; DB 15; Length 987;
Best Local Similarity 49.6%; Pred. No. 0.00017;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 80 CCAAGCGCGGAGTGTGAGGTTCGAGAGCAGAGGCTGCGGACCGCGCCCGCTCCCT 139
DB 260 CCGCGGTGCGAGCGCGCTTCGAGACCTTGACACACAGCGTCCGACCGCGCGCGCT 319

QY 140 TCCTCCCGGCTACAGACAGTCTGCGCGGAGTACGTGACAGCCACAGCGGAC 139
DB 320 TCGACACAGTCTTGGCGCGGACCTTCTTCGGTACCTGGGAGACAGCCGACCTGCT 319
QY 200 AGCTCATGCGCCCTGCTGCTCGCGATGGGTCCCGGTGCTGACAGCCGCTGACGCTGG 259
DB 380 GGCCTCTACAAAGCGCGGATGAGCCAGGACCGGTGGAAATCGCCGCTGCTGCTGCGCG 439
QY 260 TGCAGTGAACCCGCGCGACGCTGCGTGCAGTGTCTCACTGCGGCGCGACGCTGC 319
DB 440 ACCAGAGCTTCTCCGGGCTCCGACCGTGTGACGCTCGGCGGCGGAGACGCTGC 439
QY 320 AGCTGCAAGGTGCGGAGCCAGCCCC 345
DB 500 TGGCGCGGTGCTGCGCGGACCC 525

RESULT 13

US-10-084-846A-1
Sequence 1, Application US/10084846A
Publication No. US2004006026A1
GENERAL INFORMATION:
APPLICANT: WEITNAUER, GABRIELE
APPLICANT: MUHLLENWEG, AGNES
APPLICANT: TREPZER, AXEL
APPLICANT: BECHTHOLD, ANDREAS
TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
FILE REFERENCE: 1974-005
CURRENT APPLICATION NUMBER: US/10/084, 846A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: PCT/EP01/09815
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: DE 101 09 166.4
PRIOR FILING DATE: 2001-02-25
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 1
LENGTH: 59816
TYPE: DNA
ORGANISM: Streptomyces viridochromogenes
US-10-084-846A-1

Query Match 4.0%; Score 51.6; DB 15; Length 59816;
Best Local Similarity 49.6%; Pred. No. 0.0013;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 80 CCAAGCCGGGATGCTGAGAGTTTGGAGCCAGGGCTGCGACCCGCCCCCTCTCT 139
DB 45600 CCGCGGTGCGACGCGCTGGAGACCTTGACCAACGCTCCGACCGCGCGGCGCT 45659
QY 140 TCCTCCCGGCTACAGACAGTCTGCGCGGAGTACGTGACAGCCACAGCGGAC 139
DB 45660 TCGACACAGTCTTGGCGCGGACCTTCTTCGGTACCTGGGAGACAGCCGACCTGCT 45719
QY 200 AGCTCATGCGCCCTGCTGCTCGCGATGGGTCCCGGTGCTGACAGCCGCTGACGCTGG 259
DB 45720 GGCCTCTACAAAGCGCGGATGAGCCAGGACCGGTGGAAATCGCCGCTGCTGCGCG 45779
QY 260 TGCAGTGAACCCGCGCGACGCTGCGTGCAGTGTCTCACTGCGGCGCGACGCTGC 319
DB 45780 ACCAGAGCTTCTCCGGGCTCCGACCGTGTGACGCTCGGCGGCGGAGACGCTGC 45839
QY 320 AGCTGCAAGGTGCGGAGCCAGCCCC 345
DB 45840 TGGCGCGGTGCTGCGCGGACCC 45865

RESULT 14

US-10-084-846A-2/c
Sequence 2, Application US/10084846A
Publication No. US2004006026A1
GENERAL INFORMATION:
APPLICANT: WEITNAUER, GABRIELE

APPLICANT: MUHLLENWEG, AGNES
APPLICANT: TREPZER, AXEL
APPLICANT: BECHTHOLD, ANDREAS
TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
FILE REFERENCE: 1974-005
CURRENT APPLICATION NUMBER: US/10/084, 846A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: PCT/EP01/09815
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: DE 101 09 166.4
PRIOR FILING DATE: 2001-02-25
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 2
LENGTH: 59816
TYPE: DNA
ORGANISM: Streptomyces viridochromogenes
US-10-084-846A-2

Query Match 4.0%; Score 51.6; DB 15; Length 59816;
Best Local Similarity 49.6%; Pred. No. 0.0013;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 80 CCAAGCCGGGATGCTGAGAGTTTGGAGCCAGGGCTGCGACCCGCCCCCTCTCT 139
DB 14217 CCGCGGTGCGACGCGCTGGAGACCTTGACCAACGCTCCGACCGCGGCGCT 14158
QY 140 TCCTCCCGGCTACAGACAGTCTGCGCGGAGTACGTGACAGCCACAGCGGAC 139
DB 14157 TCGACACAGTCTTCCGGGCTCCGCGGAGTCTTCGCTACCTGGGAGCACCCGACCTGCT 14098
QY 200 AGCTCATGCGCCCTGCTGCTCGCGATGGGTCCCGGTGCTGACAGCCGCTGACGCTGG 259
DB 14097 GGCCTCTACAAAGCGCGGATGAGCCAGGACCGGTGGAAATCGCCGCTGCTGCGCG 14038
QY 260 TGCAGTGAACCCGCGCGACGCTGCGTGCAGTGTCTCACTGCGGCGCGACGCTGC 319
DB 14037 ACCAGAGCTTCTCCGGGCTCCGACCGTGTGACGCTCGGCGGCGGAGACGCTGC 13978
QY 320 AGCTGCAAGGTGCGGAGCCAGCCCC 345
DB 13977 TGGCGCGGTGCTGCGCGGACCC 13952

RESULT 15

US-10-156-761-6879
Sequence 6879, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARDO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 6879
LENGTH: 1185
TYPE: DNA
ORGANISM: Streptomyces avermitilis

FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1185)
US-10-156-761-6879

Query March 4.0%; Score 51; DB 14; Length 1185;
Best Local Similarity 45.0%; Pred. No. 0.00029;
Matches 230; Conservative 0; Mismatches 280; Indels 1; Gaps 1;

QY 182 ACAACCAACGAGCGGACAGCTATGACCTGCTGTGCGAGATGAGTCCCGATCGTCA 241
DB 503 ACCGCGCGCTGTGTGCGCGCCCTCCAGGAGCGGCTCCCGGCGCGGTGACCTGCGAGAACG 562
QY 242 GCACCCCTGACGCTGCGGTGAGTGAACCCGCGCGGACGCGCTGCGTCACTTTCAC 301
DB 563 AGACCAACTGTGCGCGCTGCGCGACAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 622
QY 302 TCGGCGCGCGACGCTGCAAGCTGCAAGGTGCGGACAGCCCGACGCGCGATCGGATT 361
DB 623 TCGT 682
QY 362 CTTGTCAACCCGCTGT 421
DB 683 GCGGTGCGCGCTGCGGCGGACCGCGCGAGATCGGCTTCTGCGGTACCGGATCGGCG 742
QY 422 CCGCGTCTGT 481
DB 743 CACTGCGCTGT 802
QY 482 GGCAGACACCCAGAGGAGAGAGGAGAGCGCGCATCTGTGAGACCCGAGAACCGGAGC 541
DB 803 TCGCCGCGCTGT 862
QY 542 CGAGAGAGGTGTG-CGCGAGAGAAAGCGGTGCGCGCGCGAGCGGAGAGGCGGATGTT 600
DB 863 CGGCGCGCTGT 922
QY 601 CAGGCTGACAGGCGAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 923 TCGACGTCTGT 982
QY 661 GCGGCGATGAGT 691
DB 983 CCGGCTGT 1013

Search completed: April 6, 2004, 20:47:15
Job time : 538 secs

CC gynaecological and contraceptive activity. Agents which modulate OI-180,
 CC OI-184 and OI-236 may be used to treat cell proliferative or degenerative
 CC disorders, associated with abnormal expression of these ovary specific
 CC genes. This ovary-specific sequence can be used as reagents to evaluate
 CC potential contraceptives, to block ovulation in a reversible manner. It
 CC is also used to screen for genetic mutations in signalling pathways, that
 CC are associated with some forms of human infertility or gynaecological
 CC cancers

XX Sequence 1276 BP; 279 A; 389 C; 394 G; 214 T; 0 U; 0 Other;

Query Match 99.1%; Score 1265; DB 3; Length 1276;

Best Local Similarity 99.9%; Pred. No. 9, 2e-291;

Matches 1276; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 1 AAGCGCGGCGAGGCGCGGAGCAGCAGCCAGTTCCTCCGCGAGAGCAGTTCACCCCTGCCCCG 60
DB 1 AAGCGCGGCGAGGCGCGGAGCAGCAGCCAGTTCCTCCGCGAGAGCAGTTCACCCCTGCCCCG 60
QY CATCTTATCCGAGAGCCACCAAGCGGGAGTGGCTGAGGTTGGAGAGCAAGGGCTGCG 120
DB CATCTTATCCGAGAGCCACCAAGCGGGAGTGGCTGAGGTTGGAGAGCAAGGGCTGCG 120
QY 61 CATCTTATCCGAGAGCCACCAAGCGGGAGTGGCTGAGGTTGGAGAGCAAGGGCTGCG 120
DB 61 CATCTTATCCGAGAGCCACCAAGCGGGAGTGGCTGAGGTTGGAGAGCAAGGGCTGCG 120
QY 121 CGAAGCGGCGCGCGCTTCTCTCCCGGCTACAGACGCTCATGCGCGAGTACGTC 180
DB 121 CGAAGCGGCGCGCGCTTCTCTCCCGGCTACAGACGCTCATGCGCGAGTACGTC 180
QY 121 CGAAGCGGCGCGCGCTTCTCTCCCGGCTACAGACGCTCATGCGCGAGTACGTC 180
DB 121 CGAAGCGGCGCGCGCTTCTCTCCCGGCTACAGACGCTCATGCGCGAGTACGTC 180
QY 181 GACAGCGCAGCAGGCGGAGAGCTCATGCGCGCTGCTGCGAGTGGGTCCCGGCTGTC 240
DB 181 GACAGCGCAGCAGGCGGAGAGCTCATGCGCGCTGCTGCGAGTGGGTCCCGGCTGTC 240
QY 181 GACAGCGCAGCAGGCGGAGAGCTCATGCGCGCTGCTGCGAGTGGGTCCCGGCTGTC 240
DB 181 GACAGCGCAGCAGGCGGAGAGCTCATGCGCGCTGCTGCGAGTGGGTCCCGGCTGTC 240
QY 241 AGCAGCGCTGACGCTGCGGTGACAGTGAACCCGCGCGAGAGCGCTCGGTGACGTGTC 300
DB 241 AGCAGCGCTGACGCTGCGGTGACAGTGAACCCGCGCGAGAGCGCTCGGTGACGTGTC 300
QY 241 AGCAGCGCTGACGCTGCGGTGACAGTGAACCCGCGCGAGAGCGCTCGGTGACGTGTC 300
DB 241 AGCAGCGCTGACGCTGCGGTGACAGTGAACCCGCGCGAGAGCGCTCGGTGACGTGTC 300
QY 301 CTGCGGCGCGGCGAGCGCTGACAGTGGTGGAGGTCGAGCCGCGAGTCCGATCGGT 360
DB 301 CTGCGGCGCGGCGAGCGCTGACAGTGGTGGAGGTCGAGCCGCGAGTCCGATCGGT 360
QY 301 CTGCGGCGCGGCGAGCGCTGACAGTGGTGGAGGTCGAGCCGCGAGTCCGATCGGT 360
DB 301 CTGCGGCGCGGCGAGCGCTGACAGTGGTGGAGGTCGAGCCGCGAGTCCGATCGGT 360
QY 361 TCTGTCAACCCCGTGGCGCAGCGCGGCGGAGAGTCCCGGAGTCTCTGAGACCGTA 420
DB 361 TCTGTCAACCCCGTGGCGCAGCGCGGCGGAGAGTCCCGGAGTCTCTGAGACCGTA 420
QY 361 TCTGTCAACCCCGTGGCGCAGCGCGGCGGAGAGTCCCGGAGTCTCTGAGACCGTA 420
DB 361 TCTGTCAACCCCGTGGCGCAGCGCGGCGGAGAGTCCCGGAGTCTCTGAGACCGTA 420
QY 421 GCGCGCTTCTGCTGCTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGGC 480
DB 421 GCGCGCTTCTGCTGCTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGGC 480
QY 421 GCGCGCTTCTGCTGCTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGGC 480
DB 421 GCGCGCTTCTGCTGCTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGGC 480
QY 481 AGCAGACACCCAGAGGAGAGAGGAGAGCGCGGAGTCTCTGAGGAGCCGAGACCGGAG 540
DB 481 AGCAGACACCCAGAGGAGAGAGGAGAGCGCGGAGTCTCTGAGGAGCCGAGACCGGAG 540
QY 481 AGCAGACACCCAGAGGAGAGAGGAGAGCGCGGAGTCTCTGAGGAGCCGAGACCGGAG 540
DB 481 AGCAGACACCCAGAGGAGAGAGGAGAGCGCGGAGTCTCTGAGGAGCCGAGACCGGAG 540
QY 541 CCGAGAGAGGTGGCGCGAGAGAAAGCGGTCCCGAGCGCGAGAGCGAGAGGCGATGTT 600
DB 541 CCGAGAGAGGTGGCGCGAGAGAAAGCGGTCCCGAGCGCGAGAGCGAGAGGCGATGTT 600
QY 541 CCGAGAGAGGTGGCGCGAGAGAAAGCGGTCCCGAGCGCGAGAGCGAGAGGCGATGTT 600
DB 541 CCGAGAGAGGTGGCGCGAGAGAAAGCGGTCCCGAGCGCGAGAGCGAGAGGCGATGTT 600
QY 601 CAGGCTGACAGGAGCGCGGAGTGGAGAGAGAGCCAGCCAGAGAGAGCGAGAGAGTGTG 660
DB 601 CAGGCTGACAGGAGCGCGGAGTGGAGAGAGAGCCAGCCAGAGAGAGCGAGAGAGTGTG 660
QY 601 CAGGCTGACAGGAGCGCGGAGTGGAGAGAGAGCCAGCCAGAGAGAGCGAGAGAGTGTG 660
DB 601 CAGGCTGACAGGAGCGCGGAGTGGAGAGAGAGCCAGCCAGAGAGAGCGAGAGAGTGTG 660
QY 661 GCGGCGATGAGTCTTGAAGCTTGGAGAGAGAGCCAGTGTCTGCGAGAGTGGCTCAG 720
DB 661 GCGGCGATGAGTCTTGAAGCTTGGAGAGAGAGCCAGTGTCTGCGAGAGTGGCTCAG 720
QY 661 GCGGCGATGAGTCTTGAAGCTTGGAGAGAGAGCCAGTGTCTGCGAGAGTGGCTCAG 720
DB 661 GCGGCGATGAGTCTTGAAGCTTGGAGAGAGAGCCAGTGTCTGCGAGAGTGGCTCAG 720
QY 721 GACCGCGGATGATCGGATGCGCTGAGAGCAGAGGCTCCCGCAAGAGCAGGAGAGAGAC 780
DB 721 GACCGCGGATGATCGGATGCGCTGAGAGCAGAGGCTCCCGCAAGAGCAGGAGAGAGAC 780
QY 721 GACCGCGGATGATCGGATGCGCTGAGAGCAGAGGCTCCCGCAAGAGCAGGAGAGAGAC 780
DB 721 GACCGCGGATGATCGGATGCGCTGAGAGCAGAGGCTCCCGCAAGAGCAGGAGAGAGAC 780
QY 781 AAGGAGCGCTGCTGCTTCAAGTCTTGAAGAGAGAGTACGCTACTATCATGCAAGAC 840
DB 781 AAGGAGCGCTGCTGCTTCAAGTCTTGAAGAGAGAGTACGCTACTATCATGCAAGAC 840
QY 781 AAGGAGCGCTGCTGCTTCAAGTCTTGAAGAGAGAGTACGCTACTATCATGCAAGAC 840
DB 781 AAGGAGCGCTGCTGCTTCAAGTCTTGAAGAGAGAGTACGCTACTATCATGCAAGAC 840
QY 841 TGGAAATCCGCTGGAGAGAGCGCTATGTGTGTGTGTGACAGGACCAAGTAAAGTGTTA 900
DB 841 TGGAAATCCGCTGGAGAGAGCGCTATGTGTGTGTGTGACAGGACCAAGTAAAGTGTTA 900

```

```

DB 841 TGGAAATCCGCTGGAGAGAGCGCTATGTGTGTGTGTGACAGGACCAAGTAAAGTGTTA 899
QY 901 CTTCAACAGTTCCTGCGAGTGTGTGAGAAATCTTACACCTTACAGAGTGAAGACAT 960
DB 900 CTTCAACAGTTCCTGCGAGTGTGTGAGAAATCTTACACCTTACAGAGTGAAGACAT 959
QY 961 CACTGTCAAGTGTGTAAGAAATTAATGTGCTGCGGAGTCAAGATTTGGCCACGTGA 1020
DB 960 CACTGTCAAGTGTGTAAGAAATTAATGTGCTGCGGAGTCAAGATTTGGCCACGTGA 1019
QY 1021 CCTTAAAGCGCCCGCATGCGAGACGCTGTGAGAGATGCAAGAGCAACGCGTGTCTG 1080
DB 1020 CCTTAAAGCGCCCGCATGCGAGACGCTGTGAGAGATGCAAGAGCAACGCGTGTCTG 1079
QY 1081 CGACAGACCTTTCAGCTTCAATATCATATTTATGTGAGAGTGGAAAAGTTTGTCTAGA 1140
DB 1080 CGACAGACCTTTCAGCTTCAATATCATATTTATGTGAGAGTGGAAAAGTTTGTCTAGA 1139
QY 1141 TGGGAGTAAATGAGTGAACAAGTGAAGTCTTCTCCCGCTTCACTCTTCCCTTCAAAAT 1200
DB 1140 TGGGAGTAAATGAGTGAACAAGTGAAGTCTTCTCCCGCTTCACTCTTCCCTTCAAAAT 1199
QY 1201 TCTTCATGACAGACAGTGTACTTGTGATATTAAGCTGTGAATPAAAGTATTGCAAAACA 1260
DB 1200 TCTTCATGACAGACAGTGTACTTGTGATATTAAGCTGTGAATPAAAGTATTGCAAAACA 1259
QY 1261 AAAAAAAAAAAAAAAAAA 1277
DB 1260 AAAAAAAAAAAAAAAAAA 1276

RESULT 2
ID ABZ24589 standard; cDNA; 1258 BP.
AC ABZ24589;
DT 31-MAR-2003 (first entry)
XX Mouse ovary-specific OI-180 cDNA.
XX Ovary; OI-180; mouse; contraceptive; antiinfertility; cytostatic;
XX gene therapy; gene; ss.
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 26..1108
XX FT //tag= a
XX FT //product= "Murine OI-180"
XX FT //transl_except= (pos:182..184,aa:Ser)
XX FT //transl_except= (pos:1004..1007,aa:Phe)
XX
XX W0200288314-A2.
XX
XX PD 07-NOV-2002.
XX PF 26-APR-2002; 2002W0-US013245.
XX
XX PR 27-APR-2001; 2001US-00844864.
XX
XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX PA (AMHP ) WYETH.
XX
XX PI Matzuk MM, Wang P, Bai Y, Wu X;
XX
XX DR WPI; 2003-167110/16.
XX
XX DR P-PSDB; ABP8235.
XX
XX PT New ovary-specific genes comprising OI-180 or OI-236, useful for
XX decreasing conception or enhancing fertility, or for the preparation of a
XX composition for treating e.g. cancer.

```


FT	exon	5325..5324
FT	/+tag= f	
FT	/number= 3	
FT	5325..6206	
FT	/+tag= g	
FT	/number= 3	
FT	exon	6207..6352
FT	/+tag= h	
FT	/number= 4	
XX		
XX	M0200289314-A2.	
XX		
XX	07-NOV-2002.	
PD		
XX		
XX	26-APR-2002; 2002MO-US013245.	
PX		
XX	27-APR-2001; 2001US-00844864.	
PA	(BAYU) BAYLOR COLLEGE MEDICINE,	
XX	(AMHP) WYETH.	
XX		
PI	Matzuk MM, Wang P, Bai Y, Wu X;	
DR	WPI, 2003-167110/16.	
DR	P-PADB; ABP58235.	
XX		
PT	New ovary-specific-genes comprising OI-180 or OI-236, useful for decreasing conception or enhancing fertility, or for the preparation of a composition for treating e.g. cancer.	
PS		
XX	Claim 1, Page 130-134; 14ip; English.	
CC	The present sequence is that of the murine ovary-specific OI-180 gene on chromosome 5. OI-180 clones were initially identified in a cDNA subtractive hybridisation screening using ovaries from Gdfc knockout mice and wild-type mice. OI-180 cDNA (see AB224589) was then used to isolate the OI-180 gene from a genomic library generated from mouse 129/SvEv strain. Loss of OI-180 results in female infertility and subfertility.	
CC	The invention provides ovary-specific and oocyte-specific murine and human OI-180, OI-184 and OI-236 polynucleotides and polypeptides. These genes and their protein products appear to relate to various cell proliferative or degenerative disorders, especially those involving ovarian tumours, such as germ line tumours and granulosa cell tumours, or infertility, such as premature ovarian failure. The invention provides a method for detection of a cell proliferative or degenerative disorder of the ovary, which is associated with the expression of OI-180, OI-184 or OI-236. It also provides a method for treating such disorders by using an agent which suppresses or enhances the respective activities of OI-180, OI-184 or OI-236, and a method of screening for compounds that interact and/or modulate the expression or activity of the ovary-specific genes. These compounds are possible contraceptive or fertility enhancing agents. The modulator is preferably a polypeptide, small molecule or polynucleotide sequence	
SQ	Sequence 6873 BP; 1716 A; 1624 C; 1697 G; 1836 T; 0 U; 0 Other:	
	Query Match 62.6%; Score 799.4; DB 7; Length 6873;	
	Best Local Similarity 99.9%; Pred.No. 8,1e-180;	
	Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
OY	3 GGGCGGCGAGGCCGGGAGCAGCACCATGTATCCCGGCGACGACGTTCAACCCTTCGCCGCA 62	
Dd	2513 GGGCGGCGAGGCCGGGAGCAGCACCATGTATCCCGGCGACGACGTTCAACCCTTCGCCGCA 257	
OY	63 TCCTTATCCGACAGCCACCAAAGCCGGGAGTGCTGAGGTTCCGAGCCCAAGGAGCTGCAG 122	
Dd	2573 TCCTTATCCGACAGCCACCAAAGCCGGGAGTGCTGAGGTTCCGAGCCCAAGGAGCTGCAG 2633	
OY	123 ACCCGGAGCCCCCTCTCTCTCCCGGCGTACAGACGTACTATGACCGCGGAGTAAGTACA 182	
Dd	2633 ACCCGGAGCCCCCTCTCTCTCCCGGCGTACAGACGTACTATGACCGCGGAGTAAGTACA 2691	
OY	183 CAAGCCACCAAGCGGCGACAGCTCATGAGCCTGCTGTGCGAGTAGGATCCCAGTGGTCAAG 242	

Db	2693	CAGCCACAGGGGAGACAGCTCAAGGCCCTGCTCTGCGATGAGTATCCCGGATCGGTCAAG	2755
QY	243	CAGCCGTGACCCCTCGGTGACAGTGAACCCGCGCGGACGCTCGGTGACAGTGTCACT	302
Db	2753	CAGCCGTGACGCTCGGTGACAGTGAACCCGCGCGGACGCTCGGTGACAGTGTCACT	2812
QY	303	CGGCGCCGCGACAGCTGACAGCTGACAGGATGCGAGGCCAGCCCGACGCCGATGGGTTTC	362
Db	2813	CGGCGCGCGACAGCTGACAGCTGACAGGATGCGAGGCCAGCCCGACGCCGATGGGTTTC	2872
QY	363	CTGTCAACCCCGTGAACAGCGCGCGCGCGGAGATCCCGGATCTGTGGACACGTAGC	422
Db	2873	CTGTCAACCCCGTGAACAGCGCGCGCGCGGAGATCCCGGATCTGTGGACACGTAGC	2932
QY	423	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGAGCAG	482
Db	2933	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGAGCAG	2992
QY	483	GCAGACACCCACGAAAGGAGAGGAGGCGCGGACATCCCTCGGAGACCCGCGGAACCGAGACC	542
Db	2993	GCAGACACCCACGAAAGGAGAGGAGGCGCGGACATCCCTCGGAGACCCGCGGAACCGAGACC	3052
QY	543	GAGAGAGGTGCGCGCGAGGAAAGCGGTCTCCCGACCGCGGAAGCAGAGAGGCGATGTTCA	602
Db	3053	GAGAGAGGTGCGCGCGAGGAAAGCGGTCTCCCGACCGCGGAAGCAGAGAGGCGATGTTCA	3112
QY	603	GGCTGCGAGAGGAGCGCGGTGAGAGCAGACGACACACCGGAGAGACCGGACAGGTGAGC	662
Db	3113	GGCTGCGAGAGGAGCGCGGTGAGAGCAGACGACACACCGGAGAGACCGGACAGGTGAGC	3172
QY	663	GGCGATGACAGTCTGAGCCTGCGAGACGAGAGCAGTGTCTTCCCGCAGAGATGCTCAAGA	722
Db	3173	GGCGATGACAGTCTGAGCCTGCGAGACGAGAGCAGTGTCTTCCCGCAGAGATGCTCAAGA	3232
QY	723	CCCCGGTATTCGATGCTCCCTCGAGACAGAGCTCTCCCGGAAAGCACGAGCAGAGACA	782
Db	3233	CCCCGGTATTCGATGCTCCCTCGAGACAGAGCTCTCCCGGAAAGCACGAGCAGAGACA	3292
QY	783	GGAGCGCCTGCGGTTCCAGTT 803	
Db	3293	GGAGCGCCTGCGGTTCCAGGT 3313	
RESULT 4			
ABZ24591			
ID	ABZ24591	standard; DNA; 4090 BP.	
XX	ABZ24591;		
XX	AC		
XX	31-MAR-2003	(first entry)	
XX	DT		
XX	XX		
DE	XX	Mouse ovary-specific OI-180 pseudogene.	
XX	XX		
KW	XX	Ovary; OI-180; mouse; contraceptive; antiinfertility; cytosstatic;	
XX	XX	gene therapy; pseudogene; ss.	
XX	OS	Mus musculus.	
XX	XX		
FT	Key	Location/Qualifiers	
FT	FT	1..4090	
FT	misc_feature	/*cag= a	
FT	FT	/note= "N represents unknown base"	
FT	FT	1..786	
FT	FT	/*cag= b	
FT	FT	/number= 1	
FT	FT	787..2616	
FT	FT	/*cag= c	
FT	FT	/number= 1	
FT	FT	2617..2709	
FT	FT	/*cag= d	
FT	FT	/number= 2	
FT	FT	2710..2817	
FT	FT	intron	

PS Claim 2; Page 136-137; 141pp; English.

XX The present sequence is that of a human ovary-specific O1-180
CC polynucleotide. In mice, loss of O1-180 is associated with female
CC infertility and subfertility. The invention provides ovary-specific and
CC oocyte-specific murine and human O1-180, O1-184 and O1-236
CC polynucleotides and polypeptides. These genes and their protein products
CC appear to relate to various cell proliferative or degenerative disorders,
CC especially those involving ovarian tumours, such as germ line tumours and
CC granulosa cell tumours, or infertility, such as premature ovarian
CC failure. The invention provides a method for detection of a cell
CC proliferative or degenerative disorder of the ovary, which is associated
CC with the expression of O1-180, O1-184 or O1-236. It also provides a
CC method for treating such disorders by using an agent which suppresses or
CC enhances the respective activities of O1-180, O1-184 or O1-236, and a
CC method of screening for compounds that interact and/or modulate the
CC expression or activity of the ovary-specific genes. These compounds are
CC possible contraceptive or fertility enhancing agents. The modulator is
CC preferably a polypeptide, small molecule or polynucleotide sequence
XX

SC Sequence 2075 BP; 588 A; 413 C; 473 G; 601 T; 0 U; 0 Other;

Query Match 10.5%; Score 134.6; DB 7; Length 2075;
Best Local Similarity 71.1%; Pred. No. 6.6e-22;
Matches 207; Conservative 0; Mismatches 79; Indels 5; Gaps 2;

QY 968 CAAAGTGTAAAGAAAGTATGCTGCTCCCACTGAGATTGCGACGTGACCTTAA 1027
DB 1360 CAGAGTTGTAAACAAACAGATGTTCCGCGCACTTAAACTCCGACGTCACCTTAA 1419
QY 1028 CGCCCCCATGGGCAAGCTTGTGGGAGATGACGAGCAAGACCCGCTGCGAGAGC 1087
DB 1420 CGGCCCCACCGTCAAGATTGTGCGTGATGACAAAGCCACCCCTGCTGACAGC 1479
QY 1088 ACCTTACGCTTCAATATCATCTTA-CTGAGAGTGCMAAAGCTTGTCTAGATGGGC 1146
DB 1480 ACTTCAGCTTCAATATCATCTTAAGTGAAGTGTGCTGTCATGCGCTGATG 1539
QY 1147 TAATGAGATGACAAAGTATGCTTCTCCCTTCCCTTCCCTTCCCTTCAAAATTTCTCA 1206
DB 1540 GAGTAGAGAGAGTATGCTTCTGCTGCTTCCCTTCCCTTCCCTTCAAAATTTCTCA 1599
QY 1207 TGACAGACAGTGTACTTGGATATTAAGCTGTGAATTAAGATTTGCA 1257
DB 1600 TGAAGGAGAGTGTATCTG---AAAAAGCTTCAATTAAGATTTGCA 1646

RESULT 6
AAK53491
ID AAK53491 standard; DNA; 114955 BP.

AC AAK53491;
XX
XX 05-JUL-1999 (first entry)
XX
XX Human adenosine A1 receptor antisense oligonucleotide fragment.
DE
XX Antisense oligonucleotide; multiple target; antisense treatment;
XX impaired respiration; inflammation; lung disease;
XX pulmonary vasoconstriction; inflammation; allergic rhinitis;
XX acute asthma; allergy; asthma; impeded respiration;
XX respiratory distress syndrome; pain; cystic fibrosis;
XX pulmonary hypertension; pulmonary vasoconstriction; emphysema;
XX chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
XX colon cancer; breast cancer; lung cancer; pancreatic cancer;
XX hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
XX prostate cancer; ss.
XX
XX Synthetic.
OS
XX WO913886-A1.
XX
XX PD 25-MAR-1999.

XX
PF 17-SEP-1998; 98WO-US019419.
XX
PR 17-SEP-1997; 97US-0059160P.
PR 09-JUN-1998; 98US-0003972.
XX
XX (UYEC-) UNIV EAST CAROLINA.
XX
XX Nyce JW;
XX WPI; 1999-229400/19.
XX
XX New antisense oligonucleotides used in treatment of, e.g. pulmonary
XX vasoconstriction.
PT
PS Disclosure; Page 37; 120pp; English.

XX The specification describes antisense oligonucleotides (AAK52869-X55271)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of RNAs corresponding to target genes, gene initiation
CC codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-
CC end and the juxta-section between coding and non-coding regions and all
CC segments of RNAs encoding proteins associated with one or more diseases,
CC conditions or mixtures. The antisense oligonucleotides may be derived
CC from sequences AAK5572-74. These multiple target oligonucleotides
CC (specifically AAK55180-271) can be used for the antisense treatment of
CC diseases and conditions. Typical diseases and conditions are those
CC associated with impaired respiration and inflammation, including lung
CC diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,
CC acute asthma, allergies, asthma, impeded respiration, respiratory
CC distress syndrome, pain, cystic fibrosis, pulmonary hypertension,
CC pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary
CC disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.
CC colon cancer, breast cancer, lung cancer, pancreatic cancer,
CC hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as
CC well as all types of cancers which may metastasize or have metastasized
CC to the lungs, including breast and prostate cancer
XX

XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;

Query Match 4.6%; Score 58.8; DB 2; Length 114955;
Best Local Similarity 32.6%; Pred. No. 0.0025;
Matches 251; Conservative 71; Mismatches 441; Indels 8; Gaps 3;

QY 3 GCGCGGAGAGGCGCGGAGAGACACCATGTTCCCGGAGAGAGTTCACACCTTGCCTGCGCA 62
DB 104641 GGCBCGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104700
QY 63 TCTTATTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 122
DB 104701 GC---GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104756
QY 123 ACCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
DB 104757 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104816
QY 183 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
DB 104817 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104876
QY 243 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 104877 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104936
QY 301 CTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 104937 CCBGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104996
QY 361 TCTGTCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 104997 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105056
QY 421 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

Db 105057 GGCCTGGGGCCGSSNNNDNNGCTCBGGGCGCGCCGCCGGCCCGGGCCGSSNNNDNCCBGGGCGCG 105116
 Qy 481 AGCGAGACACCCACGAGGAGAGGGAGACCCTGGCATCTTCGGGAGACCCGGAGACCGGAG 540
 Db 105117 CCGCCCGCGGGGCGSSNNNDNCCBGGGCGCGCCGC--CGGGCCGGCCGSSNNNDNBGGGC 105174
 Qy 541 CCGAGAGAGGTGCGCCGCGAGGAAGGGTCCCCCGACCGCGAGAGGAGGAGGATGTT 600
 Db 105175 GCGCCCGCGCGGGCGCCGSSNNNDNNGGCGCGCCCGCCGGCCCGSSNNNDNNGGCGCG 105234
 Qy 601 CAGGCTGCAGAGGAGCGCCGGGTGGAGACAGACCCAGACCGGAGACCGGAAAGTGTG 660
 Db 105235 CCGCCCGCGCGCGCCGSSNNNDNNGCGCGCGCCGGCCGSSNNNDNCCCGCCGCGCG 105294
 Qy 661 GCGGCGATGCACTCTAGGCTGTGGAGCGAGACCATGTCTGTGCGGAGAGATGGTCA 720
 Db 105295 CCGGCGCGSSNNNDNNGCGCGCGCGCGCGCGSSNNNDNNGCGCGCGCGCGCGSSN 105354
 Qy 721 GACCCCGGATTCGGATGCGCCCTCGAGACAGGCTCTCCGCAAGCAG 771
 Db 105355 NDNNGCGCGCGCGCGSSNNNDNCCGCGCGCGCGCGSSNNNDNNG 105405
 RESULT 7
 ID ADA71938 standard; DNA; 2000 BP.
 xx ADA71938;
 AC ADA71938;
 AT 20-NOV-2003 (first entry)
 xx DE
 xx Rice gene, SEQ ID 5263.
 KW Plant; bacterial infection; fungal infection; viral infection; rice;
 gene; ds.
 OS Oryza sativa.
 XX
 XX WO2003000898-A1.
 PN
 XX 03-JAN-2003.
 PD
 XX 22-JUN-2001; 2001WO-IB001105.
 PF
 XX 22-JUN-2001; 2001WO-IB001105.
 PR
 XX 22-JUN-2001; 2001WO-IB001105.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
 PI Karaganti P, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
 DR WPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 PS
 PS Claim 27; SEQ ID NO 5263; 893bp; English.
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 CC
 XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

	Query Match	4.5k; Score 57.8; DB 7; Length 2000;
	Best Local Similarity 10.4k; Pred. No. 0.0012;	
	Matches 79; Conservative 359; Mismatches 315; Indels 7; Gaps 4;	
Qy	519 CTGCGGAGACCCGGAGACCGGAGCCGAGAGGTGGCCCGC-AGGAAAGCGGTCCCCAGC 577	
Db	12 CSMRGRRBRYTMAGMMSCAMGSSRMSRKMGSMTXKXKSSCGCKXTTRRKSMTYASS 71	
Qy	578 CGGCAACGAGAGAGGGGATGTTCAGGCTCAGGAGCGAGCCGGGTGGAGCAGCAGCAC 637	
Db	72 ASGRITGKMSGSGSYSGMGKXKRKRSRMTGRGRGRGRBRBRMTMGVYRRCARSGRMAAG 131	
Qy	638 CACCGAGAGACCCGGAACAGTGTGCGGGCGAT-GCAGTCTAGCCTGTGAGCGAGAGCCA 696	
Db	132 SGRMMWGKSRMSYMMWCYARGCSCKRKKSGSGWGTCKRRGARGGSGWSGATYKXGSM 191	
Qy	697 TGTCTTCGCGAGAGATAGTGTCAAGAACCCCGGTGATTCCGATGCGCCCTCGAGACCAAGCC 756	
Db	192 SKRMMMSCGSGCGRSAYSRYYTGRSKTGYTKMTYASRCMRMYMTTSYMASSY 251	
Qy	757 TCCCCGGAAAGCAGGAGCAGAGCAAGAGACGCGCTCGTTCCAGTTCTTAGACAGAAAG 816	
Db	252 TWCSRKRSMWTKMTRMRMSRSGMTSWSYKMMCTAYKXKSYSPMCYWRGGMGAT 311	
Qy	817 TAACGGTACATCATCTGCAGAGCTGTGAATCCGGTGGGAGAGCGGCTATGTGTGGTG 876	
Db	312 RYWRGWSRDAWMTYKMTYRGYKMGKGMWAGMMRMSCRSKACIYWRMMDMTTR 371	
Qy	877 GTGACAGGCAACAGTAGTGTTACTTCAACAGATTGCGAGGTGTGAGAAATCCTA 936	
Db	372 RRRWAKKSRTSRKXKKKCMCRKRKYRMRGYSR--RRSCRARMMRRCRSGRAMKXGCR 429	
Qy	937 CAACCTTACAGATGAGAGACATCACCTGTCAAAGTTGTAAAAAGAACTGATGTGCCTG 996	
Db	430 GCMTCRMKSYGMKWKMSKMWASKTKMMSRMTYRRKKCSRTTWMGTRGSMGT--MG 486	
Qy	997 CCCAGTCAGATTGCGCACGTGACCCCTAAACGCCCCCATGCGCAAGACTTGTGTGGAG 1056	
Db	487 RCRYKXKSGMKRCKRRBRMGMYMRMKRYMSARITMYRCARKYGSYSAARKRACRWYRG 546	
Qy	1057 ATGCAAGAGCAAAAGCCTGTCTCCGACAGACACTTGAAGCTTCAATAACATCATTTAGT 1116	
Db	547 KGYWAGKMMRMYRMWYKMMWYKRYKSKSYSCMSYTCMSYTASCMKSHRAGAKCMRSK 606	
Qy	1117 AGAGTCGAAAACGTTTCTGTAGATGGGGCTAATGAATGAGCAAGTAGAGCTTCTCCC 1176	
Db	607 MSASKMSRBSRKRCKAKSKRSSAKRYAMMGMTSGSRMSMKSYTCYWRKMGSMKSYCT 666	
Qy	1177 TCTTCACCTCTCTCCCTTCCAAATTTCTCATGACAGACAGTGTAAGTATATAAGCC 1236	
Db	667 WMYYSKTYAKYKSYWRYRYRAACMTWRYRYRYSYMTYAMYSITSMATGMYKIS 726	
Qy	1237 TGTGAATTAAGGTATTGCAACCAAAAAAATTTTTAAAA 1275	
Db	727 GRWTSWYXKCGSMKRYRSMWYYSMMWAKTWMKMBRYA 765	
RESULT 8		
ADA69587		
ID	ADA69587 standard; DNA; 1032 BP.	
XX		
AC	ADA69587;	
DT	20-NOV-2003 (first entry)	
XX		
DE	Rice gene, SEQ ID 2910.	
XX		
KW	Plant; bacterial infection; fungal infection; viral infection; rice;	
XX	gene; ds.	
OS	Oryza sativa.	
XX		

PN W02003000898-A1.
 XX 03-JAN-2003.
 XX 22-JUN-2001; 2001MO-IB001105.
 PF 22-JUN-2001; 2001MO-IB001105.
 PR 22-JUN-2001; 2001MO-IB001105.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Kategriri F, Quan S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
 DR WPI; 2003-175290/17.
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 6; SEQ ID NO 2910; 899pp; English.
 PS The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX Sequence 1032 BP; 189 A; 331 C; 394 G; 118 T; 0 U; 0 Other;
 SQ
 Query Match 4.2%; Score 54.2; DB 7; Length 1032;
 Best Local Similarity 43.4%; Pred. No. 0.007;
 Matches 251, Conservative 0; Mismatches 328; Indels 0; Gaps 0;
 QY 137 CTTTCTCCCGGTAGACAGCTCATGCGCCGCGAGTACGTCAGACCCACGCGG 196
 DB 167 CTTTCTTACACACCCCGACAGCTCTCTGAAAGAGATCTAGAGAGCTCCCG 226
 QY 197 CACAGCTATGCGCTCTGCTGCGGATGATGCTCCCGGTGCTGACAGCCGTCG 256
 DB 227 AGAAGAGCGCGCTGACCGCGAGAGAGTGCATGCTGAGAGAGAGCTTCGAG 286
 QY 257 CGGTGAGTGAACCCGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 316
 DB 287 AGAACAAGCTGAGACCGGAGCGAGAGAGAGCTGCGCGCGAGAGCTGCGAG 346
 QY 317 TGACGCTGAGAGGTGCGGAGCGAGAGCGCGAGCTGCTGCTGCTGCTGCTG 376
 DB 347 GCGAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 406
 QY 377 GCGACGCGCGCGCGGAGATCCCGGATCTGCGAGACCGTTCGCTGCTGCTG 436
 DB 407 GCGACTTGAACCGCTCAAGCGCGCTTGAACCGCTTCCCGCGAGAGAGCTTC 466
 QY 437 TGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496
 DB 467 TCCAGGAAACACCGCTCTCACTGATGATGATGATGATGATGATGATGATG 526
 QY 497 AGGAGAGAGGAGCGCGCATCTCTGCGGAGACCGGAGACCGGAGAGAGTGGCG 556
 DB 527 AGGAGAGAGAGCGCGAGAGGAGCGCGCGCGCGCGCTGAGTCTGCGCGG 586
 QY 557 CGAAGAAAGGCTGCTGCGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
 DB 587 CGGCGAG 646
 QY 617 CCGGTGAG 676

DB 647 CGGCTTGAAG 706
 QY 677 AGCTGAG 715
 DB 707 CGGCGGAG 745
 RESULT 9
 ID AAX53491 standard; DNA; 114955 BP.
 XX AAX53491;
 AC AAX53491;
 XX 05-JUN-1999 (first entry)
 DT
 XX Human adenosine A1 receptor antisense oligonucleotide fragment.
 DE Antisense oligonucleotide; multiple target; antisense treatment;
 XX impaired respiration; inflammation; lung disease;
 XX pulmonary vasoconstriction; inflammation; allergic rhinitis;
 XX acute asthma; allergy; asthma; impeded respiration;
 XX respiratory distress syndrome; pain; cystic fibrosis;
 XX pulmonary hypertension; pulmonary vasoconstriction; emphysema;
 XX chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
 XX colon cancer; breast cancer; lung cancer; pancreatic cancer;
 XX hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
 XX prostate cancer; ss.
 XX Synthetic.
 OS W09913886-A1.
 XX 25-MAR-1999.
 PD 17-SEP-1998; 98WO-US019419.
 PF 17-SEP-1997; 97US-0059160P.
 PR 09-JUN-1998; 98US-00093972.
 XX (UNEC-) UNIV EAST CAROLINA.
 PA Nyce JW;
 PI WPI; 1999-223400/19.
 DR New antisense oligonucleotides used in treatment of, e.g. pulmonary
 PT vasoconstriction.
 PS Disclosure; Page 37; 120pp; English.
 XX The specification describes antisense oligonucleotides (AAX52869-X55271)
 CC directed against at least 2 mRNAs selected from target genes, coding and
 CC non-coding regions of RNAs corresponding to target genes, gene initiation
 CC codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-
 CC end and the juxta-section between coding and non-coding regions and all
 CC segments of RNAs encoding proteins associated with one or more diseases,
 CC conditions or mixtures. The antisense oligonucleotides may be derived
 CC from sequences AAX55180-271. These multiple target oligonucleotides
 CC (specifically AAX55180-271) can be used for the antisense treatment of
 CC diseases and conditions. Typical diseases and conditions are those
 CC associated with impaired respiration and inflammation, including lung
 CC diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,
 CC acute asthma, allergies, asthma, impeded respiration, respiratory
 CC distress syndrome, pain, cystic fibrosis, pulmonary hypertension,
 CC pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary
 CC disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.
 CC colon cancer, breast cancer, lung cancer, pancreatic cancer,
 CC hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as
 CC well as all types of cancers which may metastasize or have metastasized
 CC to the lungs, including breast and prostate cancer
 XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;
 SQ

Query Match 4.1%; Score 52; DB 2; Length 114955;
 Best Local Similarity 32.3%; Pred. No. 0.1;
 Matches 244; Conservative 50; Mismatches 456; Indels 5; Gaps 2;

```

QY      13 GCGGGGAGACGACCACTGTTCCCGGAGACGACGTTCCACCCCTGCGCCGACCTCTTACCG 72
DB      11059 GCGGAGGAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 111000
QY      73 CAGGCGACCAAGCGCGGAGATGCTGAGAGTTGAGAGCTGAGAGCTGACCGCGCC 132
DB      110999 GCGGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 110940
QY      133 CCTCTCTTCTCTCC---GGCTACAGACAGCTCATGCGCGGAGTACTGCAACGCA 188
DB      110939 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110880
QY      189 CCAAGCGGACAGACTCATGCGCTGCTGCGGATGCGGATGCGCGGATGCGCGGATGCGCGG 248
DB      110879 NSGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110820
QY      249 TGACGCTGCGGTGAGGTGAACCCGCG-CCGCGAGCGCTCGGTGAGTTCAGTCTCGGCG 307
DB      110819 CNHNNNNCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110760
QY      308 GCGCGACGCTGACGCTGAGAGGTGCGGAGTCCGAGCGCGCGCGCGCGCGCGCGCGCGCG 367
DB      110759 GACGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110700
QY      368 AACCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 427
DB      110699 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110640
QY      428 TCTGCTGCTGACCTTCTGTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487
DB      110639 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110580
QY      488 CACCGACGAGGAGAGGAGGAGCGCGCGGATCTCGGAGAACCGGAGCGGAGAG 547
DB      110579 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110520
QY      548 AGGTGCGCGCGAGGAGAAAGCGGTCTCCCGACCGCGAGCGAGGAGGAGGATGTTCAAGCT 607
DB      110519 NHHNNNSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110460
QY      608 CAGGCGCAGCGCGGCTGGAGAGCAGACCCACGAGGAGACCGGAAAGTGTGGCGCGA 667
DB      110459 CGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110400
QY      668 TGCACTGTGAGCTGTGAGAGGAGGAGGATGCTCTGCGCGAGAGATGCTCAAGACCCCG 727
DB      110399 GCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110340
QY      728 GTGATTGAGTGTGCTCTGAGACCAAGGCTCTCCCG 762
DB      110339 NHHNNNSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110305

```

RESULT 10

AB237562
 ID AB237562 standard; DNA, 987 BP.

```

AC      AB237562;
XX
XX      26-FEB-2003 (first entry)
XX
DE      Streptomyces viridochromogenes Av1G4 encoding polynucleotide.
XX
XX      Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
XX      medicine; Staphylococcus aureus; biosynthetic gene cluster; gene; ds.
XX
OS      Streptomyces viridochromogenes.
XX
PN      WO200268436-A1.

```

```

XX      06-SEP-2002.
PD
XX      24-AUG-2001; 2001WO-EP009815.
XX
XX      25-FEB-2001; 2001DE-01009166.
XX
XX      (COMB-) COMBINATURE BIOPHARM AG.
XX
XX      Wettenauer G, Muehlenweg A, Trefzer A, Bechtold A;
XX      WPI: 2003-018650/01.
XX      P-FSDB; ABP76728.
XX
XX      New avilamycin derivatives, useful for treatment of infections, and
XX      PT nucleic acid encoding avilamycin synthesis enzymes.
XX
XX      Claim 13; Page 244-248; 319pp; German.
XX
XX      The invention relates to avilamycin derivatives (I) with antibacterial,
XX      CC virucide, protozoacide and fungicide activity. (I) are useful for
XX      CC treatment of infections (bacterial, viral, protozoal or fungal), in human
XX      CC or veterinary medicine, particularly where caused by Staphylococcus
XX      CC aureus. (I) are more hydrophilic than known avilamycins. The present
XX      CC sequence is that of a nucleic acid encoding avilamycin synthesis enzymes
XX      CC from the Streptomyces viridochromogenes Avilamycin A biosynthetic gene
XX      CC cluster (AB237515-AB237516)
XX
SQ      Sequence 987 BP; 121 A; 358 C; 371 G; 137 T; 0 U; 0 Other;

```

Query Match 4.0%; Score 51.6; DB 7; Length 987;
 Best Local Similarity 49.6%; Pred. No. 0.029;
 Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```

QY      80 CCAAGCGCGGAGTACCTGAGAGTTGAGACCGAGGAGCTGCGGACCGCGCGCGCGCTCT 139
DB      260 CCGCGGTGCGCGAGCGCGCTGTGACGACCTGACACAGAGGTGCGGACCGCGCGCGCGCT 319
QY      140 TCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTGACAGCAGCAGCGGCGAC 199
DB      320 TGACACAGCTCTTCCGCGGAGACTTCTTGGCTACTGCGGAGACACCCGACCTGTGTGT 379
QY      200 AGCTCATGCGCTGCTGTGCGGAGTGTCCCGGTGCGTCAAGACCGCTGACGCTGCGG 259
DB      380 GGCTTACAAAGCGGCGGAGTAGACGAGGACACGAGTGGATGCGCGGCTGTGTGCGGCGC 439
QY      260 TCGAGTGAACCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
DB      440 ACCAGGACTTCTCCGCGGCTGCGACCGTGTGAGCTGCGGAGCGGAGACGAGACGCTGC 499
QY      320 AGCTGCAAGGTGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345
DB      500 TGGCGCGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 525

```

RESULT 11

AB237516/c
 ID AB237516 standard; DNA, 59816 BP.

```

AC      AB237516;
XX
XX      26-FEB-2003 (first entry)
XX
DE      Streptomyces viridochromogenes Av1 gene cluster sense strand.
XX
XX      Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
XX      medicine; Staphylococcus aureus; biosynthetic gene cluster; gene; ds.
XX
OS      Streptomyces viridochromogenes.
XX
PN      WO200268436-A1.
XX
XX      06-SEP-2002.

```

```
XX 24-AUG-2001; 2001MO-EP009815.
PF 25-FEB-2001; 2001DE-01009166.
XX (COMB-) COMBINATURE BIOPHARM AG.
XX (COMB-) COMBINATURE BIOPHARM AG.
XX Weitenauer G, Muehlenweg A, Trefzer A, Bechtold A;
XX WPI; 2003-018650/01.
DR P-PSDB; ABP76680, ABP76681, ABP76682.
XX New avilamycin derivatives, useful for treatment of infections, and
PT nucleic acid encoding avilamycin synthesis enzymes.
XX
XX Example 1; Page 68-301; 31pp; German.
XX The invention relates to avilamycin derivatives (I) with antibacterial,
CC virucide, protozoacide and fungicide activity. (I) are useful for
CC treatment of infections (bacterial, viral, protozoal or fungal), in human
CC or veterinary medicine, particularly where caused by Staphylococcus
CC aureus. (I) are more hydrophilic than known avilamycins. The present
CC sequence is that of a nucleic acid encoding avilamycin synthesis enzymes
CC from the Streptomyces viridochromogenes Avilamycin A biosynthetic gene
CC cluster (AB237515-AB237516)
XX
SQ Sequence 59816 BP; 8760 A; 21053 C; 21088 G; 8915 T; 0 U; 0 Other;
Query Match 4.0%; Score 51.6; DB 7; Length 59816;
Best Local Similarity 49.6%; Pred. No. 0.11;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 80 CCAAGCGCGGGAGTGGCTGAGGTTGAGCCAGGCGCTGCCGCCGCCCTCTCT 139
DB 14217 CCGCGGTGCGAGCGCTGAGCCAGCTGAGCCAGCGCTGCCGCCGCCCTCT 14158
QY 140 TCTTCCCGGCTACAGACGCTCATGCGCGCGGAGTACGTGACACGCGCGGAC 199
DB 14157 TCGACACGCTTTCGGCGGAGCTTCTTCGCGTACCTGGGAGACGCCGACCTGTCT 14098
QY 200 AGCTCATGCGCTCTGCTGTCGCGATGAGTCCCGGCTGCTGACAGCCGCTGACG 259
DB 14097 GGCTCTACACGCGCGGATGAGCCAGGCGACCGGTGGATCGCGGCTGTGCGGCGC 14038
QY 260 TGCAGGTGAACCCGCGCGGACGCTCGGTGACGCTTCACCTCGGCGCGCGCGCTGC 319
DB 14037 ACCAGGACTTCTCGGCGTCCGACGCTGTGACGTGCGGCGGCGGAGACGCTGC 13978
QY 320 AGCCTGACGGGTGCCGACGACGCC 345
DB 13977 TGGCGCGGTGCTGCGCGGCGCACCCC 13952
RESULT 12
AB237515
ID AB237515 standard; DNA; 59816 BP.
XX
XX AB237515;
AC 05-MAR-2003 (first entry)
XX
XX Streptomyces viridochromogenes Avi gene cluster sense strand.
DE Streptomyces viridochromogenes Avi gene cluster sense strand.
XX Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
KW Avilamycin; Staphylococcus aureus; biosynthetic gene cluster; gene; ds.
XX
XX Streptomyces viridochromogenes.
OS Streptomyces viridochromogenes.
XX
XX WO200268436-A1.
XX
XX 06-SEP-2002.
XX
XX 24-AUG-2001; 2001MO-EP009815.
PF
```

```
XX 25-FEB-2001; 2001DE-01009166.
XX (COMB-) COMBINATURE BIOPHARM AG.
XX (COMB-) COMBINATURE BIOPHARM AG.
XX Weitenauer G, Muehlenweg A, Trefzer A, Bechtold A;
XX WPI; 2003-018650/01.
DR P-PSDB; ABP76678, ABP76679.
XX New avilamycin derivatives, useful for treatment of infections, and
PT nucleic acid encoding avilamycin synthesis enzymes.
XX
XX Example 1; Page 68-301; 31pp; German.
XX The invention relates to avilamycin derivatives (I) with antibacterial,
CC virucide, protozoacide and fungicide activity. (I) are useful for
CC treatment of infections (bacterial, viral, protozoal or fungal), in human
CC or veterinary medicine, particularly where caused by Staphylococcus
CC aureus. (I) are more hydrophilic than known avilamycins. The present
CC sequence is that of a nucleic acid encoding avilamycin synthesis enzymes
CC from the Streptomyces viridochromogenes Avilamycin A biosynthetic gene
CC cluster (AB237515-AB237516)
XX
SQ Sequence 59816 BP; 8915 A; 21088 C; 21053 G; 8760 T; 0 U; 0 Other;
Query Match 4.0%; Score 51.6; DB 7; Length 59816;
Best Local Similarity 49.6%; Pred. No. 0.11;
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 80 CCAAGCGCGGGAGTGGCTGAGGTTGAGCCAGGCGCTGCCGCCGCCCTCTCT 139
DB 45600 CCGCGGTGCGAGCGCTGAGCCAGCTGAGCCAGCGCTGCCGCCGCCCTCT 45659
QY 140 TCTTCCCGGCTACAGACGCTCATGCGCGCGGAGTACGTGACACGCGCGGAC 199
DB 45660 TCGACACGCTTTCGGCGGAGCTTCTTCGCGTACCTGGGAGACGCCGACCTGTCT 45719
QY 200 AGCTCATGCGCTCTGCTGTCGCGATGAGTCCCGGCTGCTGACAGCCGCTGACG 259
DB 45720 GGCTCTACACGCGCGGATGAGCCAGGCGACCGGTGGATCGCGGCTGTGCGGCGC 45779
QY 260 TGCAGGTGAACCCGCGCGGACGCTCGGTGACGCTTCACCTCGGCGCGCGCTGC 319
DB 45780 ACCAGGACTTTCGGCGGTCCGACGCTGTGAGACGTGCGGCGGCGGAGACGCTGC 45839
QY 320 AGCCTGACGGGTGCCGACGACGCC 345
DB 45840 TGGCGCGGTGCTGCGCGGCGCACCCC 45865
RESULT 13
AAV4436
ID AAV4436 standard; DNA; 400 BP.
XX
XX AAV4436;
AC 17-OCT-2003 (revised)
XX
XX 09-NOV-1998 (first entry)
XX
XX Mycobacterium tuberculosis antigen XP22 5' DNA.
DE Mycobacterium tuberculosis antigen XP22 5' DNA.
XX Tuberculosis; infection; diagnosis; antigen; XP22; ss.
KW Mycobacterium tuberculosis; strain Erdman.
XX
XX WO9816645-A2.
XX
XX 23-APR-1998.
XX
XX 07-OCT-1997; 97WO-US018214.
XX
XX 11-OCT-1996; 96US-00729622.
PF
```

PR 13-MAR-1997; 97US-00818111.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ;
XX
XX WPI; 1998-251292/22.
XX
XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used to
PT develop products for the detection of M. tuberculosis infection and
PT diagnosis of tuberculosis.
XX
XX
XX Claim 4; Page 185-186; 250pp; English.
XX
XX This is the 5' region of DNA coding for an antigenic portion of
CC Mycobacterium tuberculosis antigen XP22; 3' DNA is provided in AAV4437.
CC XP22 DNA was isolated from a M. tuberculosis strain Erdman genomic DNA
CC expression library using sera from patients having extrapulmonary
CC tuberculosis. It bears no similarity to known sequences. The invention
CC relates to methods and compositions for diagnosing tuberculosis. It
CC provides polypeptides (see AAV64291-W64379) comprising an antigenic
CC portion of a soluble M. tuberculosis antigen, or an immunogenic portion
CC of a M. tuberculosis antigen, as well as DNA sequences encoding such
CC polypeptides, recombinant expression vectors and transformed or
CC transfected host cells. Also claimed are methods and diagnostic kits for
CC detecting M. tuberculosis infection in a patient using these
CC polypeptides, antibodies or oligonucleotide probes and primers. (updated
CC on 17-Oct-2003 to standardise OS field)
XX
XX
SQ Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 U; 0 Other;
Query Match 3.9%; Score 49.6; DB 2; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.064;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 458 CCTCACTGAGAGTTGCGGAGGAGGAGACCCAGAAAGGAGGAGCCCGGCAT 517
DB 119 CTTACCCCAAGGCGCCGAGCAACGCGGCAACGGGTGAGCGGGGTCCGGCGCA 178
QY 518 CTTGCGGAGACCCGAGAACCGGAGCGAGAGGTGCGCGAGAAAGCGGTCCCGCAGC 577
DB 179 CGGCGGAAAGCGCGGAAACCGCGGAGCAACACACACCGCGCGCGGACACAG 238
QY 578 CGGCAACCGAGAGAGGCGATGTTGAGCTGAGGCGAGCGCGGTGGAGCAACCA 637
DB 239 CGGCGAGCGGCGGCGCGCGCGCGGAGACCGGCGGAAACCGGCGAACCCCGGCA 298
QY 638 CACCGAGAGACCGGAGACAGTGTGGCGGAGTGCAGTGTGAGCTGGAGAGGAGCCAT 697
DB 299 CGGCAACCGGCGGCGGCGGCGGCGGAGACCGGCGGAAACCGGCGAACCCCGGCA 358
QY 698 GTCTGCGGAGAGATGCTCAGG 721
DB 359 CGGCGGCGAGCGGTGCACTCTCAGG 382
RESULT 14
AAV64545
ID AAV64545 standard; DNA; 400 BP.
XX
XX AAV64545;
XX
XX 27-JAN-1999 (first entry)
XX
XX M. tuberculosis immunogenic polypeptide XP22 5'-end DNA.
XX
XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
XX vaccine; pharmaceutical; infection; diagnosis; ss.
XX
XX Mycobacterium tuberculosis.
XX
XX WO9816646-A2.
XX
XX

XX
XX 23-APR-1998.
PD
XX
XX 07-OCT-1997; 97WO-US018293.
PF
XX
XX 11-OCT-1996; 96US-00730510.
PR
XX
XX 13-MAR-1997; 97US-00818112.
XX
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ;
XX
XX WPI; 1998-261042/23.
XX
XX
XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
PT develop products for the detection of M. tuberculosis infection and for
PT diagnosis, treatment and prevention of tuberculosis.
XX
XX
XX Claim 4; Page 175; 230pp; English.
XX
XX This sequence encodes an immunogenic portion of a soluble Mycobacterium
CC tuberculosis (MT) antigen which can be used in a method for inducing
CC protective immunity against tuberculosis (TB). This sequence can be
CC formulated into vaccines and/or pharmaceutical compositions for
CC immunising against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis
XX
XX
SQ Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 U; 0 Other;
Query Match 3.9%; Score 49.6; DB 2; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.064;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 458 CCTCACTGAGAGTTGCGGAGGAGGAGACCCAGAAAGGAGGAGCCCGGCAT 517
DB 119 CTTACCCCAAGGCGCCGAGCAACGCGGCAACGGGTGAGCGGGGTCCGGCGCA 178
QY 518 CTTGCGGAGACCCGAGAACCGGAGCGAGAGGTGCGCGAGAAAGCGGTCCCGCAGC 577
DB 179 CGGCGGAAAGCGCGGAAACCGCGGAGCAACACACACCGCGCGCGGACACAG 238
QY 578 CGGCAACCGAGAGAGGCGATGTTGAGCTGAGGCGAGCGCGGTGGAGCAACCA 637
DB 239 CGGCGAGCGGCGGCGCGCGCGGAGACCGGCGGAAACCGGCGAACCCCGGCA 298
QY 638 CACCGAGAGACCGGAGACAGTGTGGCGGAGTGCAGTGTGAGCTGGAGAGGAGCCAT 697
DB 299 CGGCAACCGGCGGCGGCGGCGGCGGAGACCGGCGGAAACCGGCGAACCCCGGCA 358
QY 698 GTCTGCGGAGAGATGCTCAGG 721
DB 359 CGGCGGCGAGCGGTGCACTCTCAGG 382
RESULT 15
AAZ19134
ID AAZ19134 standard; DNA; 400 BP.
XX
XX AAZ19134;
XX
XX 05-NOV-1999 (first entry)
XX
XX M. tuberculosis recombinant antigen DNA encoding 5' XP22.
XX
XX Antigen; diagnosis; detection; infection; antibody; immunisation;
XX vaccine; immunity; ss.
XX
XX Mycobacterium tuberculosis.
XX
XX WO9942118-A2.
XX
XX 26-AUG-1999.
XX
XX

XX 17-FEB-1999; 99WO-US003265.
 XX 18-FEB-1998; 98US-00024753.
 PR 05-MAY-1998; 98US-00072596.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R,
 PI Vedavick TS, Twardzik DR, Lodes MT, Hendrickson RC,
 XX WPI; 1999-527416/44.
 DR
 XX
 XX New polypeptide comprising antigenic portions of M. tuberculosis.
 PT
 XX
 PS Claim 4; Page 215-216; 323pp; English.
 XX
 CC This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against M.
 CC tuberculosis infection. The new detection methods are needed as current
 CC vaccination strategies do not provide 100% immunity
 XX
 SQ Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 U; 0 Other;

Query Match 3.9%; Score 49.6; DB 2; Length 400;
 Best Local Similarity 49.2%; Pred. No. 0.064;
 Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY	458	CCCTCCTGAGAGGTTGGGGAGGAGACACCAAGGAGAGGGAGCCCGGCAT	517
DB	119	CTTCACCCAGAGGCGCCAGCAGCCGCGAAGCGGCTGACGCGGGTGGCGGCA	178
QY	518	CCTCGGGGAGCCCGGAGACCGAGAGGTTGCGGAGAAAGCGTCCCCAGC	577
DB	179	CGGCGGAAAGCGGGAACGGCGCAGACACACACCGCGCGCGCGGACCAAGG	238
QY	578	CGGAGAGCGAGAGGGGATGTTCAAGCTGCAAGGCGAGCGGGTGGAGCAGCCAC	637
DB	239	CGGCGAGCGGCGGGCGCGCGGCGGAGCCGCGGAGCCGCGCGCAGC	298
QY	638	CACCGAGAGACCGGAAACAGTGTGCGCGGATGCACTGAGCCTGGAGCGAGAGCCAT	697
DB	299	CGGACACCGGCGGCGCAGACAGCAGCGGCGGCAACCGCGGCGGCAAGGCGGCGC	358
QY	698	GTCCTGCCGAGAGATGGCTCAGG	721
DB	359	CGGCGGCGAGCGTGCCTCTCAGG	382

Search completed: April 6, 2004, 18:08:48
 Job time : 622 secs